



**Isabel Maria da Silva
Machado**

**A Internet como Factor de Inovação no Acesso a
Informação Estatística**



**Isabel Maria da Silva
Machado**

**A Internet como Factor de Inovação no Acesso a
Informação Estatística**

dissertação apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Gestão da Ciência, Tecnologia e Inovação, realizada sob a orientação científica do Professor Doutor Carlos Manuel dos Santos Ferreira, Professor Associado do Departamento de Economia, Gestão e Engenharia Industrial da Universidade de Aveiro e sob a co-orientação da Professora Doutora Maria Paula de Brito Duarte Silva, Professora Associada da Faculdade de Economia da Universidade do Porto.

o júri

presidente

Doutor Jorge de Carvalho Alves
Professor catedrático da Universidade de Aveiro

vogais

Doutor Carlos Manuel dos Santos Ferreira
Professor associado da Universidade de Aveiro (Orientador)

Doutora Maria Paula de Pinho de Brito Duarte Silva
Professora associada da Faculdade de Economia da Universidade do Porto (Co-Orientadora)

Doutor Paulo Jorge Mota de Pinho Gomes
Professor catedrático convidado do ISEGI – Instituto Superior de Estatística e Gestão de Informação da Universidade Nova de Lisboa

agradecimentos

Queria expressar o meu agradecimento a todos aqueles que me acompanharem no percurso desta investigação. Ao meu orientador, Professor Doutor Carlos Ferreira, por me ter despertado o interesse pela Estatística no final da minha licenciatura e assim tornar possível esta investigação. À Professora Doutora Paula Brito, não só pela sua disponibilidade e orientação, mas também pela objectividade, dedicação e rigor sempre presentes. E, em especial, ao Dr. Pedro Campos pela sua disponibilidade e contribuição valiosa no arranque deste estudo.

Por fim, agradeço a todos os amigos, família e colegas por terem suportado as minhas ausências, bem como as minhas presenças!

palavras-chave

Internet, informação estatística, inovação, pesquisa, divulgação.

resumo

A divulgação de informação através da Internet tem tido um crescimento notável a nível mundial nos últimos anos, constituindo por isso um factor de inovação no acesso a informação, em particular, a informação estatística. Tendo como base as perspectivas e estratégias empresariais ao nível da inovação, os comportamentos humanos na pesquisa de informação e respectivos modelos, e o papel da Internet na divulgação da informação estatística, esta investigação pretende avaliar a utilização deste tipo de informação (difundida através da Internet) como suporte à gestão empresarial. Para o efeito, foi efectuada uma análise de Institutos estrangeiros de Estatística, nomeadamente o *Statistics Netherlands* e *Statistics Canada*, e um estudo de caso do Instituto Nacional de Estatística português. Em todos os casos foi feita uma avaliação dos serviços prestados através dos respectivos *websites* e posterior comparação entre os serviços STATLINE, CANSIM e INFOLINE prestados pelos institutos holandeses, canadiano e português, respectivamente.

Relativamente ao caso português recorreu-se ainda à implementação de um inquérito junto de empresas potenciais utilizadoras do serviço INFOLINE do INE. Os resultados sugerem que as empresas recorrem à Internet para recolher informação estatística sobre possíveis consumidores, sobre os seus concorrentes e sobre os mercados, no sentido de formularem/modificarem e apoiarem as suas estratégias de gestão.

keywords

Internet, statistical information, innovation, search, disclosure.

abstract

Information disclosure through the Internet has shown a remarkable and global increase in the last years, so that it is now a factor of innovation to access information, in particular, statistical information. Based on management perspectives and strategies concerning innovation, human behaviour as concerns information search and the respective models, and the Internet role in statistical information disclosure, this research aims at analysing the utilization of this kind of information (spread through the Internet) to support business management. Therefore, analyses of foreign National Institutes of Statistics, namely Statistics Netherlands and Statistics Canada, and a case study of the Portuguese National Statistical Institute (INE), have been made. In all cases, the services offered through their respective websites have been evaluated. A comparison between STATLINE, CANSIM and INFOLINE services, offered, respectively, by the Dutch, Canadian and Portuguese Institutes, has also been made.

As concerns the Portuguese case, a survey on companies, potential users of the INFOLINE service, has been implemented. The results suggest that the companies use the Internet to gather statistical information about potential consumers, about their competitors and about the markets, in order to formulate/modify and support their management strategies.

Índice

INTRODUÇÃO	15
CAPÍTULO I – INOVAÇÃO: PERSPECTIVAS, SISTEMÁTICA E ESTRATÉGIA EMPRESARIAL.....	19
I.1 INTRODUÇÃO	19
I.2 O CONCEITO DE INOVAÇÃO	19
I.2.1 Fontes de Inovação.....	21
I.2.2 Tipologia da Inovação	24
I.2.3 Inovação como Factor de Competitividade	25
I.3 INOVAÇÃO EM PORTUGAL.....	26
I.4 CONCLUSÃO.....	29
CAPÍTULO II – PESQUISA E UTILIZAÇÃO DE INFORMAÇÃO	33
II.1 INTRODUÇÃO	33
II.2 COMPORTAMENTOS HUMANOS NA PESQUISA DE INFORMAÇÃO.....	33
II.2.1 Necessidades de Informação	34
II.2.2. Pesquisa de Informação	36
II.2.3 Uso da Informação.....	38
II.2.4 Um Modelo Integrado de Pesquisa de Informação	40
II.3 CONCLUSÕES	42
CAPÍTULO III – INTERNET E INFORMAÇÃO	45
III.1 A INTERNET E O SEU DESENVOLVIMENTO	45
III.2 MODELOS DE PESQUISA DE INFORMAÇÃO NA <i>WEB</i>	47
III.3 MODELO DE ELLIS DE PESQUISA DE INFORMAÇÃO	48
III.4 MODELO DE COMPORTAMENTO DE PESQUISA DE INFORMAÇÃO NA <i>WEB</i> (CHOO, DETLOR & TURNBULL, 2000).....	53
III.5 CONCLUSÃO.....	57
CAPÍTULO IV – A INTERNET COMO FONTE DE INFORMAÇÃO ESTATÍSTICA.....	61
IV.1 INTRODUÇÃO.....	61
IV.2 INFORMAÇÃO ESTATÍSTICA ANTES DA INTERNET	62

IV.3 INFORMAÇÃO ESTATÍSTICA NA INTERNET.....	64
IV.4 CONFIDENCIALIDADE/SEGREDO ESTATÍSTICO	68
IV.4.1 Segredo Estatístico Português	70
IV.5 LEVANTAMENTO DE SÍTIOS <i>WEB</i>	72
IV.5.1 Caso <i>Statistics Netherlands</i> – StatLine	73
IV.5.1.1 Serviços Disponíveis	75
IV.5.1.2 Serviço <i>StatLine</i>	76
IV.5.2 Caso <i>Statistics Canada</i> – CANSIM	82
IV.5.2.1 Serviços Disponíveis	83
IV.5.2.2 Serviço CANSIM.....	85
IV.5.3 Caso Português	91
IV.5.3.1 Instituto Nacional de Estatística (INE).....	91
IV.5.3.1.1 Serviços do INE.....	93
IV.5.3.1.2 Serviço INFOLINE do INE.....	96
IV.5.3.2 Banco de Portugal.....	103
IV.5.3.2.1 Serviços ao nível de Informação Estatística	104
IV.6 CONCLUSÃO	109
 CAPÍTULO V – NOVAS PERSPECTIVAS NA APRESENTAÇÃO/DIFUSÃO DE ESTATÍSTICAS OFICIAIS	117
V.1 INTRODUÇÃO.....	117
V.2 DADOS SIMBÓLICOS	119
V.3 ANÁLISE DE DADOS SIMBÓLICOS NOS INSTITUTOS NACIONAIS DE ESTATÍSTICA	121
V.4 <i>SOFTWARE SODAS</i>	128
V.5 ESTUDOS JÁ EFECTUADOS	129
V.6 CONCLUSÃO	131
 CAPÍTULO VI – ESTUDO DE UM CASO: SERVIÇO INFOLINE DO INE	133
PARTE A – METODOLOGIA.....	133
VI.1 INTRODUÇÃO.....	133
VI.2 INQUÉRITO	133
VI.3 CARACTERIZAÇÃO DA AMOSTRA	136
 PARTE B – ANÁLISE E INTERPRETAÇÃO DE RESULTADOS	138
VI.1 INTRODUÇÃO.....	138
VI.2 QUESTIONÁRIO REFERENTE AOS UTILIZADORES DO SERVIÇO INFOLINE DO INE	138

VI.2.1 Conhecimento do Serviço INFOLINE.....	138
VI.2.2 Frequência de Utilização do Serviço INFOLINE	139
VI.2.3 Tipo de Informação Acedida no Serviço INFOLINE.....	140
VI.2.4 Temas consultados no Serviço INFOLINE.....	143
VI.2.5 Apreciação do Serviço INFOLINE	146
VI.2.5.1 Classificação da Apreciação	148
VI.2.6 Informação Disponibilizada no Serviço INFOLINE.....	151
VI.2.7 Análise Multivariada	152
VI.3 QUESTIONÁRIO REFERENTE AOS UTILIZADORES E NÃO UTILIZADORES DO SERVIÇO INFOLINE DO INE	155
VI.3.1 Acesso a outros <i>Sítes</i> de Divulgação de Informação Estatística	155
VI.3.1.1 <i>Sítes</i> de Divulgação de Informação Estatística a que as Empresas Acedem ..	158
VI.3.1.2 Motivo de Não Utilização de <i>Sítes</i> de Divulgação de Informação Estatística...	160
VI.4. CONCLUSÕES	162
CONCLUSÕES, LIMITAÇÕES E PERSPECTIVAS	167
BIBLIOGRAFIA.....	171
REFERÊNCIAS ELECTRÓNICAS	179
ANEXOS.....	183
ANEXO I – CRITÉRIOS DE AVALIAÇÃO DE BASES DE DADOS DISPONÍVEIS <i>ONLINE</i> (COOKE, 1999).....	185
ANEXO II – QUESTIONÁRIO.....	189

Índice de Figuras

Figura 1 – Conceitos de Criatividade, Invenção e Inovação. Fonte: (Oliveira 1999).....	20
Figura 2 – Fontes de Informação para as Empresas Industriais (gráfico superior) e de Serviços (gráfico inferior). Fonte: (Conceição & Ávila 2001).	27
Figura 3 – Innovation DNA. Fonte: (InnovationNetwork 2002).	31
Figura 4 – Factores que afectam as necessidades de informação. Fonte: (Choo 2000). ..	36
Figura 5 – Factores que afectam a pesquisa de informação. Fonte: (Choo 2000).	36
Figura 6 – Pesquisa de Informação: Acesso e Qualidade. Fonte: (Choo 2001).	38
Figura 7 – Uso de Informação. Fonte: (Choo 2001).	39
Figura 8 – Modelo integrado de uso de informação. Fonte: (Choo 1998).	41
Figura 9 – Modelo de Pesquisa de Informação de Ellis.	52
Figura 10 – <i>Homepage de Statistics Netherlands</i>	74
Figura 11 – Contador <i>Online</i> de População na Holanda	76
Figura 12 – Janela principal do programa <i>Webselector – StatLine</i>	77
Figura 13 – Janela do programa <i>Webselector</i> com os sub-temas e metainformação activada	78
Figura 14 – Janela do programa <i>Webselector</i> com as variáveis específicas ao sub-tema escolhido.....	78
Figura 15 – Janela do Programa <i>Webselector</i> com a opção do período temporal	79
Figura 16 – Exemplo de quadro dinâmico gerado pelo programa <i>Webselector – StatLine</i>	79
Figura 17 – Exemplo de um gráfico gerado no <i>StatLine</i>	81
Figura 18 – Exemplo de um mapa estatístico gerado no <i>StatLine</i>	82
Figura 19 – <i>Homepage de Statistics Canada</i>	83
Figura 20 – Lista dos temas principais disponíveis na <i>CANSIM</i> , através da pesquisa por tema.....	87
Figura 21 – Pesquisa por palavra-chave na <i>CANSIM</i>	88
Figura 22 – Directório de tabelas da <i>CANSIM</i>	88
Figura 23 – Directório dos estudos e programas estatísticos realizados pelo <i>Statistics Canada</i>	89
Figura 24 – Exemplo de um sumário de uma tabela na base de dados <i>CANSIM</i>	89
Figura 25 – Exemplo de uma tabela obtida a partir da base de dados <i>CANSIM</i>	90
Figura 26 – Exemplo de um gráfico obtido a partir da base de dados <i>CANSIM</i>	91
Figura 27 – <i>Homepage</i> do Instituto Nacional de Estatística.....	93

Figura 28 – Formulário de entrada para utilizadores registados no INFOLINE.....	97
Figura 29 – Janela INFOLINE de acesso aos Indicadores Estatísticos Gerais.....	97
Figura 30 – Janela INFOLINE de acesso aos Quadros Estatísticos	98
Figura 31 – Processo de obtenção de uma Série Cronológica no Serviço INFOLINE.....	99
Figura 32 – Janela de acesso à pesquisa por Unidade Territorial do serviço INFOLINE	100
Figura 33 – Janela de acesso às Folhas de Informação Rápida do serviço INFOLINE..	100
Figura 34 – Janela de acesso aos estudos no serviço INFOLINE	101
Figura 35 – Janela de acesso ao Perfil Indicadores Estruturais do Serviço INFOLINE ..	101
Figura 36 – Janela de acesso ao Perfil Género do Serviço INFOLINE	102
Figura 37 – Programa QSS que permite a consulta, impressão e exportação dos dados que figuram nos Quadros de Situação Sectoriais, Banco de Portugal	107
Figura 38 – Exemplo de um Quadro de Situação Sectorial obtido através do programa QSS	108
Figura 39 – Exemplo de microdados e respectivos dados agregados	122
Figura 40 – Tabela simbólica com intervalos de confiança (Fonte: Marcelo, 2004).....	123
Figura 41 – Exemplo de junção de dois estudos independentes através da análise de dados simbólicos; representação por <i>zoom-star</i> (Fonte: Marcelo, 2004).....	124
Figura 42 – Processo de controlo de qualidade na produção de informação estatística	125
Figura 43 – Processo de rotação da amostra (Fonte: Marcelo, 2004)	126
Figura 44 – Tabela clássica e tabela simbólica em que os objectos simbólicos representam as partes da amostra na 1ª e 2ª rotações (Fonte: Marcelo, 2004)	127
Figura 45 – Ambiente de trabalho do software SODAS, versão 1.2	128
Figura 46 – Visão geral do <i>software</i> SODAS2 (Fontes: Bock & Diday, 2000; ASSO Developers, 2005)	129
Figura 47 – Exemplo de uma tabela simbólica no <i>software</i> SODAS. (Fonte: Exemplos SODAS)	129
Figura 48 – Diagramas de Barras dos Aspectos de Avaliação	148
Figura 49 – Dendograma segundo o Método Ward	149
Figura 50 – Primeiro Plano factorial da Análise Factorial de Correspondências Múltiplas das variáveis Apreciação do Serviço e Frequência de Utilização (Ex.: a4=5 significa aspecto 4 com pontuação 5)	153
Figura 51 – Dendograma da análise hierárquica ascendente dos inquiridos.....	154
Figura 52 – Representação da partição em três classes no primeiro plano factorial da Análise Factorial de Correspondências Múltiplas dos indivíduos	154

Índice de Tabelas

Tabela 1 – Fontes de Inovação. Fonte: (Drucker 1996).	22
Tabela 2 – Actividades de Pesquisa de Informação. Fonte: (Choo et al. 2000).....	51
Tabela 3 – Quatro modos de pesquisa de informação. Fonte: (Choo, Detlor & Turnbull, 2000).....	55
Tabela 4 – Modos e Actividades de Pesquisa de Informação na <i>Web</i>	56
Tabela 5 – Sugestões para optimizar a utilização da <i>Web</i> . Fonte: (Choo, 2000).	59
Tabela 6 – Top 10 de Fontes de Informação Estatística segundo Paula Berinstein. Fonte: (Berinstein, 1998)	66
Tabela 7 – Pontos fortes (↗) e fracos (↘) das bases de dados StatLine, CANSIM e INFOLINE segundo alguns critérios (Cooke, 1999).....	115
Tabela 8 – Tabela do inquérito sobre a ocupação do tempo no país Basco.....	119
Tabela 9 – Tabela Simbólica para uma variável multi-valores (Fonte: Bock & Diday, 2000)	120

Índice de Quadros

Quadro 1 – Empresas que utilizam computador, empresas que utilizam Internet e empresas que possuem <i>website</i> , segundo o número de pessoas ao serviço (%). Fonte: (INE 2003).	47
Quadro 2 – Caracterização da amostra.....	136
Quadro 3 – Tabela de Contingência: Sector de Actividade vs a Utilização do Serviço INFOLINE do INE (TS – Total por Sector de Actividade, TU – Total por Utilização do Serviço, T – Total)	137
Quadro 4 – Frequência de Utilização do Serviço INFOLINE.....	139
Quadro 5 – Tipo de Informação acedido no serviço INFOLINE	140
Quadro 6 – Tabela de Contingência: Sector de Actividade vs Tipo de Informação Acedido	141
Quadro 7 – Temas consultados no serviço INFOLINE.....	143
Quadro 8 – Tabela de Contingência: Sector de Actividade vs Tema Consultado.....	144
Quadro 9 – Frequências Relativas e Relativas Acumuladas para a Apreciação do Serviço INFOLINE	146

Quadro 10 – Coeficiente de Correlação de Pearson entre as variáveis relativas à questão 7	149
Quadro 11 – Encontra a informação que procura no serviço INFOLINE	151
Quadro 12 – Tabela de Contingência: Utilização do Serviço vs Acesso a outros <i>Sítes</i> de Informação Estatística	156
Quadro 13 – Teste de Independência do Qui-Quadrado para a relação Utilizador do Serviço INFOLINE/Acesso a outros <i>Sítes</i> de Informação Estatística (SPSS)	156
Quadro 14 – Tabela de Contingência: Sector de Actividade vs Acesso a outros <i>Sítes</i> de Informação Estatística (TS – Total por Sector de Actividade, TU – Total por Utilização do Serviço, T – Total)	157
Quadro 15 – Outros <i>Sítes</i> de Informação Estatística Acedidos	158
Quadro 16 – Tabela de Contingência: Sector de Actividade vs Outros <i>Sítes</i> acedidos (TS – Total por Sector de Actividade, TU – Total por Utilização do Serviço, T – Total)	159
Quadro 17 – Motivos de Não Utilização de <i>sítes</i> de Informação Estatística	160
Quadro 18 – Tabela de Contingência: Utilizador INFOLINE vs Motivos de Não Utilização de <i>sítes</i> de Informação Estatística (TM – Total por Motivo, TU – Total por Utilização do Serviço, T – Total)	161
Quadro 19 – Tabela de Contingência: Sector de Actividade vs Motivos de Não Utilização de <i>sítes</i> de Informação Estatística	162

Índice de Gráficos

Gráfico 1 – Conhecimento do Serviço INFOLINE	138
Gráfico 2 – Diagramas de Extremos e Quartis das variáveis que constituem a avaliação do Serviço INFOLINE	147
Gráfico 3 – Acesso a outros <i>Sítes</i> de Divulgação de Informação Estatística	155

Introdução

A informação e o conhecimento sempre foram elementos estratégicos no desenvolvimento da humanidade, tendo sido factores relevantes em todas as revoluções tecnológicas, e assumindo, nesta última, o carácter de revolução da própria informação (Alonso, 2002). Embora a importância da informação seja reconhecida por diferentes áreas do conhecimento, a sua definição ainda não é concretamente delimitada. Em geral, cada área de interesse designa a informação conforme a sua possível utilização. Contudo, é consenso entre essas áreas que a informação promove acréscimo de conhecimento.

Num mundo em constante “devir”, as decisões empresariais não podem ser simplesmente baseadas em factores como a intuição ou a experiência. É facto assente, que qualquer estratégia de gestão terá de ser fundamentada em informações estatísticas e não em meros ideais ou métodos fechados às mutações do mercado. Assim, a informação estatística tem como principal atributo diagnosticar o ambiente nacional e internacional para servir de suporte à formulação de políticas públicas e estudos socio-económicos, às pesquisas académicas em diversos níveis e, de modo mais abrangente, possibilitar à sociedade a construção de uma cidadania colectiva (Gracioso, 2003). Esta mais valia reconhecida à informação estatística, era aliás descrita por Fernando Pessoa em 1926 quando escrevia “Nada revela mais uma incapacidade fundamental para o exercício do comércio que o hábito de concluir o que os outros querem, sem estudar os outros (...)” in Revista de Comércio e Contabilidade.

A informação estatística possui características que lhe são exclusivas. A adjectivação estatística dada à informação elimina o carácter abstracto da informação propriamente dita, tornando-a objecto concreto de estudo. Assim, existem vários conceitos para delimitação do termo “informação estatística”, tais como (Gracioso 2003):

- ➔ Dado – consiste numa referência não elaborada, algo não interpretado, não classificado, não estruturado, não ajustado a um contexto.
- ➔ Estatística Oficial – consiste num conjunto de dados sociais, demográficos e económicos recolhidos, compilados e disponibilizados regularmente pelos institutos ligadas ao planeamento do governo e outras instituições públicas para a sociedade civil, governo e empresas.
- ➔ Indicador Social – consiste numa medida quantitativa (em geral) dotada de significado social e usada para substituir, quantificar ou operacionalizar um

conceito social abstracto, de interesse teórico (para pesquisa académica) ou programático (para formulação de políticas).

- ➔ Informação Institucionalizada – informação à qual é atribuída uma potencialidade de utilidade informacional.
- ➔ Conhecimento – conhecimento registado ou descrito que adquire um novo estatuto de informação que, por sua vez, ao ser absorvida por algum indivíduo gera um novo conhecimento.

Com o crescimento da Internet e o aumento do interesse em recolher dados estatísticos, tem crescido também o desejo de aceder a informação estatística produzida por instituições, por investigadores universitários e, em particular, pelos Institutos Nacionais de Estatística. A tarefa que se propõe desenvolver com este trabalho é a de conhecer a situação actual, ao nível de divulgação de informação estatística através da Internet, e perceber a importância da Internet como factor de inovação no acesso a este tipo de informação por parte das empresas.

Sendo o objecto de estudo a inovação e o acesso à informação estatística induzida pelo uso da Internet em contexto empresarial, irá procurar-se identificar de que modo os Institutos Nacionais de Estatística permitem este acesso e de que formas é que o utilizador pode pesquisar e visualizar, *online*, a informação estatística produzida. Identificar-se-ão Institutos estrangeiros de Estatística, nomeadamente o *Statistics Netherlands* e *Statistics Canada* e, também o Instituto Nacional de Estatística de Portugal e estudar-se-ão os serviços prestados através das suas páginas na Internet. Estabelecer-se-á ainda comparações entre os serviços STATLINE, CANSIM, e INFOLINE prestados pelos institutos holandeses, canadiano e português, respectivamente.

Quivy & Campenhoudt (1998) sugerem que o investigador deve procurar enunciar o projecto de investigação na forma de uma pergunta de partida, através da qual se tenta exprimir o que se procura saber melhor. Assim, a pergunta de partida desta investigação “será a Internet um factor de inovação no acesso a informação estatística?” reúne em si os propósitos deste trabalho.

Para o desenvolvimento desta investigação, os temas gerais a analisar são:

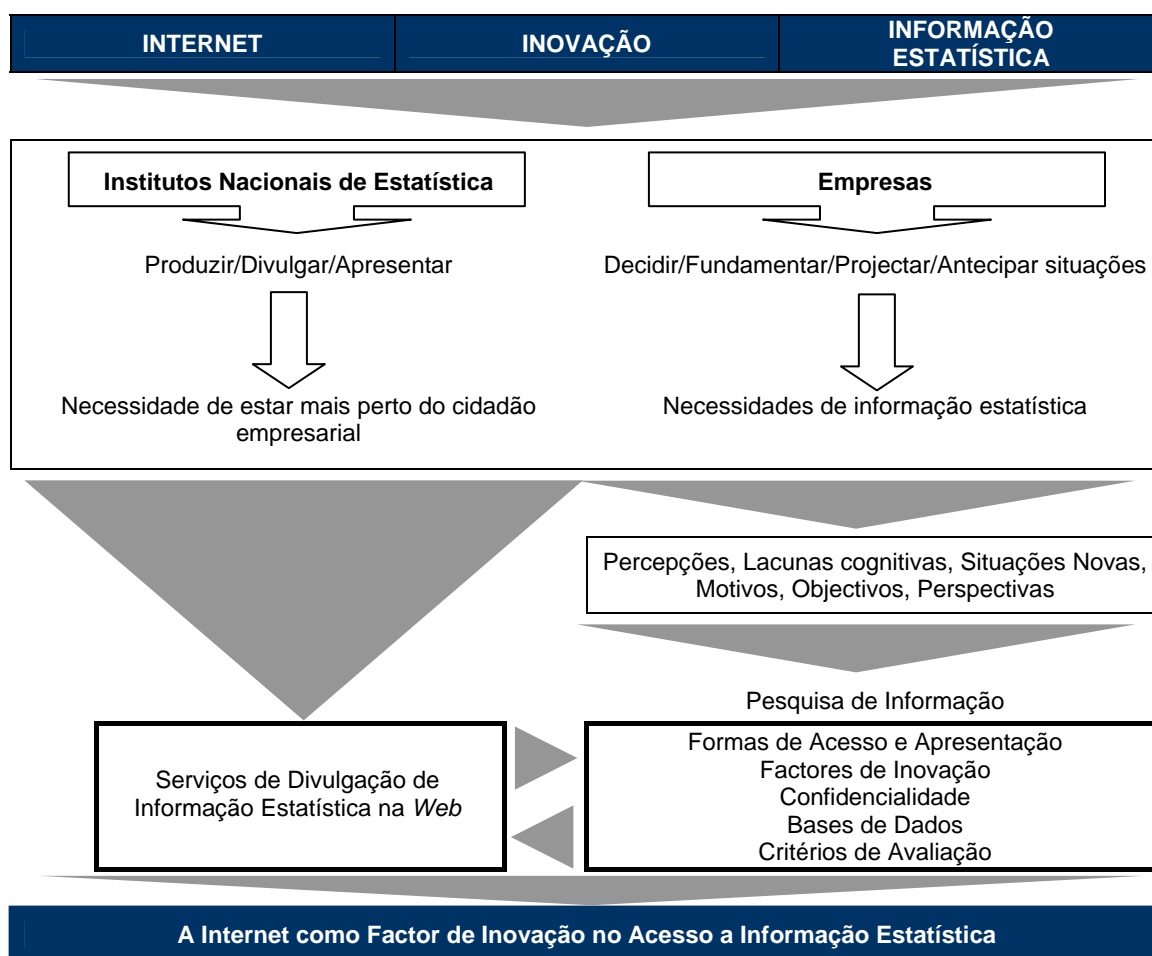
- ➔ O comportamento humano na pesquisa e utilização de informação na *web*;
- ➔ A importância da Internet no acesso a informação estatística;
- ➔ As formas de acesso e apresentação de informação estatística na Internet;
- ➔ O papel dos Institutos Nacionais de Estatística na divulgação de informação estatística;

- As novas perspectivas de apresentação e divulgação de informação estatística;
- A identificação do grau de utilização de *sítes* de divulgação de informação estatística no contexto empresarial;
- A identificação do grau de satisfação dos utilizadores de um serviço de divulgação de informação estatística.

Segundo os autores Quivy & Campenhoudt (1998), o modelo de análise é o prolongamento natural da problemática, articulando de forma operacional os marcos e as pistas para orientação da investigação. A formulação da pergunta de partida constitui uma base para a construção da hipótese geral que define o fio condutor da investigação:

A Internet constitui um acesso inovador a informação estatística e de divulgação da mesma, de modo a dar resposta às necessidades da população empresarial no que se refere à tomada de decisões a esse nível.

Na linha de pensamento da hipótese, é possível construir o modelo de análise para a investigação que se pretende realizar:



No que respeita à organização, esta dissertação está estruturada em seis capítulos para além da introdução e conclusão. O primeiro Capítulo apresenta os conceitos de inovação, perspectivas, sistemática e estratégia empresarial. No segundo Capítulo, estudam-se os comportamentos humanos na pesquisa de informação. O terceiro Capítulo aborda os modelos de pesquisa de informação na Internet. No quarto Capítulo pretende-se estudar a relação existente entre a Internet e a informação, onde são analisados os serviços prestados por Institutos Nacionais de Estatística. O quinto Capítulo pretende ilustrar uma nova perspectiva na apresentação e divulgação de informação estatística. O Capítulo VI apresenta os objectivos, a metodologia e as conclusões do estudo de caso realizado. Por fim, apresentam-se as conclusões a que se chegou.

De forma a tornar a dissertação mais completa, são ainda incluídos dois anexos:

Anexo I – apresenta uma lista de critérios de avaliação de bases de dados disponíveis *online* (Cooke, 1999) que servirá de suporte à avaliação e comparação dos serviços STATLINE, CANSIM, e INFOLINE prestados pelos Institutos Nacionais de Estatística holandês, canadiano e português, respectivamente.

Anexo II – contém o questionário utilizado no estudo de caso do serviço INFOLINE do Instituto Nacional de Estatística (INE).

Capítulo I – Inovação: Perspectivas, Sistemática e Estratégia Empresarial

I.1 Introdução

“Surgiu uma ideia que vai revolucionar os negócios”, disse Henry Benedict para Philo Remington¹, sobre a máquina de escrever *Sholes*, em meados do século XIX. Não demorou muito tempo para que a máquina de escrever se tornasse um dispositivo de presença constante no local de trabalho. Ela provocou grandes ondas de mudança no ambiente social do escritório, criando uma nova classe de funcionários para cuidar de documentos escritos. O que ocorreu com a inovação da máquina de escrever, ocorreu também com outros objectos que pareciam impossíveis de criar, mas que ao longo dos anos foram surgindo.

A inovação é um tema muito globalizante e, por isso, irão abordar-se alguns aspectos que se consideram mais relevantes para o tema em estudo. Começa-se por caracterizar o conceito de inovação e tentar perceber como é que esta nova personagem entrou e se desenvolveu no mundo da gestão empresarial.

I.2 O Conceito de Inovação

Quando se ouve a palavra “inovação”, é natural que as pessoas associem o termo ao último modelo de máquinas fotográficas digitais que viram dias antes anunciado numa revista ou ao telemóvel com capacidades admiráveis de obtenção de informação, consequentes dos avanços mais recentes no campo das telecomunicações. De facto, no dia-a-dia e no vocabulário de senso comum, a palavra inovação é ainda predominantemente relacionada com a variedade de produtos ou serviços que criam novos desenvolvimentos tecnológicos. Contudo, o conceito de inovação não se refere apenas à vertente de desenvolvimento tecnológico, mas também possui outras vertentes. Segundo Peter Drucker (1986), a inovação é “a ferramenta específica dos empresários, o meio através do qual eles exploram a mudança como oportunidade para um negócio ou um serviço diferente. É possível apresentá-la sob a forma de disciplina, aprendê-la e praticá-la”. Assim, o conceito de inovação é ampliado, não se limitando ao produto e respectiva evolução tecnológica. Pode-se falar de inovação ao nível da estratégia (Hamel e Prahalad, 1995), da gestão de recursos, da concepção ou acompanhamento dos

¹ Utterback, James M., (1996), Dominando a Dinâmica da Inovação, Brasil, QualityMark Editora, pp. 1-24.

processos, das formas de organização e estruturas da vertente financeira, da produção, da distribuição, do *marketing* e comercialização, das marcas, das políticas de remuneração e recompensa, da gestão de qualidade ou ambiental, em suma, em todas as actividades relacionadas com a forma de ser e de estar de uma organização.

Drucker (1986) considera que a inovação é tanto conceptual, como perceptual. Um dos imperativos da inovação é, pois, sair e observar, perguntar, ouvir. Consultar números e consultar pessoas. Definir analiticamente qual terá de ser o conteúdo da inovação para que ela satisfaça uma oportunidade. Depois disso, os gestores devem observar os consumidores, para verem quais são os seus valores, as suas expectativas e as suas necessidades.

Existe, também, uma abordagem à inovação, que a considera como condicionada ao juízo do mercado, ou seja dos clientes. Por exemplo, segundo Porter (1990) a inovação é “uma nova forma de fazer as coisas que é comercializada”; também Afuah (1998) refere-se à inovação como “o uso de novo conhecimento para oferecer um produto ou serviço que os clientes querem”.

Jorge Oliveira (1999) considera que o conceito de inovação é bastante mais rico e abrangente do que os conceitos de criatividade ou invenção. Contudo, os dois últimos activam a inovação. Este autor define inovação como sendo “a transformação de ideias e/ou utilização de invenções, de que resultam aplicações úteis conducentes a melhoramentos” (Figura 1).

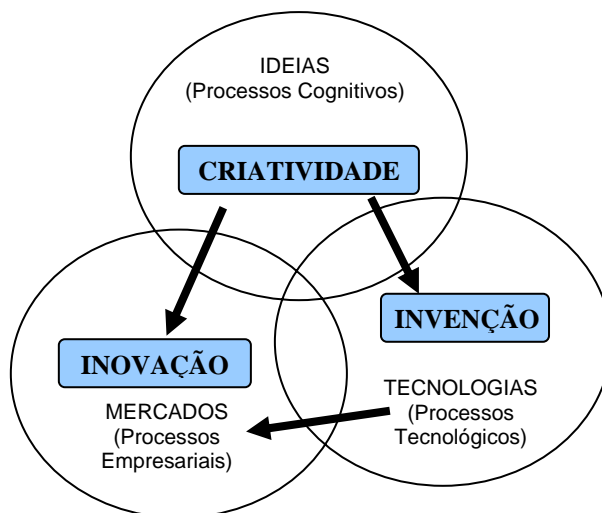


Figura 1 – Conceitos de Criatividade, Invenção e Inovação. Fonte: (Oliveira 1999).

Segundo Oliveira (1999), a criatividade é o produto do génio humano, enquanto gerador de novas ideias, conceitos ou teorias e a invenção é um passo em frente, no qual se delineia um produto, processo ou protótipo resultante da combinação de ideias. A

criatividade existe no universo de ideias, em que os processos são cognitivos; a invenção, no universo das tecnologias, em que os processos são tecnológicos; e a inovação, no universo dos mercados, em que os processos são empresariais (Figura 1).

Através da inovação (Adriano Freire, 2000), a empresa constrói no presente as bases do seu desenvolvimento futuro. De facto, são as actuais iniciativas inovadoras que dão origem aos novos produtos, serviços e processos da organização, sustentando assim o seu crescimento a médio e longo prazo. Segundo Freire (2000), a criatividade promove o espírito inventivo, que por sua vez gera inovações de sucesso para o mercado.

O Dicionário de Marketing (2001) define inovação como um dos processos empresariais mais críticos, relacionado com o acréscimo permanente de novos elementos aos produtos ou serviços. Inovação não é a criação ou invenção constante de novos produtos; é antes um processo contínuo de introdução de melhorias em produtos actuais com vista à conquista de vantagem competitiva.

Os fundadores de *Innovation Network*² (2003) criaram o [Innovation Proverbs Poster](http://www.thinksmart.com/library/index.html)³ onde apresentam uma perspectiva global sobre o que é inovação. Segundo estes autores, a inovação é: acompanhar a onda da mudança, criando futuro; encontrar uma agulha num palheiro de ideias; olhar para o mundo como se ele fosse novo; uma aplicação estruturada da imaginação; a implementação de ideias que criem valor; o resultado da liberdade ou a reacção a uma opressão; um método e uma magia.

I.2.1 Fontes de Inovação

Para Peter Drucker (1986), a inovação é o instrumento específico da actividade empresarial. Este autor considera que a inovação não tem de ser apenas técnica mas que deve assumir um carácter económico e social. Os empresários tentam criar valor e contribuir com algo, independentemente da sua motivação individual. Normalmente visam obter melhoramentos e modificações que permitam atingir, para além de novos e diferentes valores, satisfações novas e diferentes. As oportunidades são sempre criadas pela mudança. Drucker (1986) considera que a inovação pode ser sistemática: “consiste na procura organizada e intencional de mudanças, e na análise sistemática das oportunidades que tais mudanças podem proporcionar à inovação económica e social”. Estas oportunidades para criar inovação, têm como impulso sete fontes que funcionam como indicadores de mudanças que possam surgir e se encontram divididas em duas categorias (Tabela 1).

² *Innovation Network*: <http://www.thinksmart.com/index.html> (consultado em 10/09/2004).

³ *Innovation Proverbs Poster*: <http://thinksmart.com/library/index.html> (consultado em 10/09/2004).

Analisando as fontes de inovação que Drucker (1986) refere, considera-se mais importante aprofundar a inovação baseada em novos conhecimentos, por ser aquela que melhor se adequa ao tema em estudo. Segundo este autor, a inovação baseada em novos conhecimentos é a vedeta da actividade empresarial. Esta inovação é aquela que tem um período de gestação mais longo. Existe um longo espaço de tempo entre o aparecimento de novos conhecimentos e a altura em que eles passam a poder aplicar-se à tecnologia. E depois há um outro longo período antes de a nova tecnologia se transformar em produtos, processos de fabrico ou serviços de mercado. Uma outra importante característica desta inovação é o facto de nunca se basear num único factor, mas sim na convergência de vários tipos de conhecimentos diferentes, nem todos científicos ou tecnológicos.

Fontes no interior da organização que são visíveis aos olhos das pessoas	
↳	O inesperado – o êxito inesperado, o fracasso inesperado, o acontecimento externo inesperado.
↳	A incongruência – entre a realidade como realmente se apresenta e a realidade como se presume que ela seja ou como “deveria ser”.
↳	Inovação baseada em necessidades operativas.
↳	Mudanças na estrutura da indústria ou na estrutura do mercado, que apanham as pessoas desprevenidas.
Fontes de oportunidades para a inovação que implicam mudanças exteriores à organização	
↳	Factores demográficos – mudanças na população.
↳	Mudanças de percepção, atitude e significado.
↳	Novos conhecimentos científicos e não científicos.

Tabela 1 – Fontes de Inovação. Fonte: (Drucker 1996).

As características da inovação baseada em novos conhecimentos conferem-lhe requisitos específicos. Em primeiro lugar, esta requer uma cuidadosa análise de todos os factores necessários, sejam os conhecimentos em si, ou os factores sociais, económicos ou perceptíveis. A análise deverá identificar quais os factores que ainda não estão disponíveis, de modo a que os empresários decidam se estes factores que faltam podem ser criados, ou se será melhor adiarem a inovação por ela ainda não ser praticável. Um outro requisito é a definição clara de uma área de actuação na posição estratégica. A inovação não deve ser introduzida por tentativas, pelo facto de que uma inovação por si só pode provocar excitação e atrair as pessoas, o que implica que o inovador terá de acertar à primeira. Por último, o inovador que se baseia em novos conhecimentos precisa de aprender e praticar gestão empresarial.

Basicamente, Drucker (1986) estabeleceu três áreas de actuação para as inovações com base em novos conhecimentos. Em primeiro lugar, há a área de actuação que consiste no desenvolvimento de um sistema completo; a segunda consiste na definição do mercado; e a terceira área de actuação consiste em ocupar uma posição estratégica, concentrando-se num objectivo.

A inovação é considerada um elemento chave no sucesso dos negócios, facto que se tornou ainda mais evidente na sociedade pós-capitalista, a que Peter Drucker (1993) chama de sociedade baseada no conhecimento. Repare-se que o conhecimento, enquanto factor de produção, é bastante diferente dos “inputs” tradicionais, como capital e trabalho; o valor do conhecimento é incerto, e o seu potencial diferente entre os vários agentes. Numa descrição sintética, poderemos dizer que as mudanças constantes produzem situações de desequilíbrio, as quais dão origem a novas oportunidades. Esta concepção de mudança/inovação é defendida, nomeadamente, por Brazeal e Herbert (1996). Por outro lado Johannessen *et al.* (2001) defendem que a mudança que é intrínseca à inovação, deve cumprir determinados requisitos, nomeadamente, ser nova e original. Ou seja, na sua opinião o sucesso da inovação é mais determinado pela sua adopção do que pela sua natureza ou origem.

Ao longo das últimas décadas (Baptista, 1999), as empresas têm desenvolvido novos mecanismos e novas práticas que lhes facilitam uma melhor adaptação às mutações que ocorrem no ambiente em que estão inseridas. Dessa multiplicidade de novas práticas, novas ferramentas e novos conceitos progressivamente interiorizados nas várias áreas das empresas, destaca-se a Internet. A Internet é uma “ferramenta” ao dispor das empresas que suporta um conjunto de serviços, entre os quais se destacam o correio electrónico, a transferência electrónica de dados e a *World Wide Web (WWW)*⁴. Estes serviços constituem um auxiliar útil às empresas ao nível da comunicação e acesso à informação, ou seja, uma porta de entrada a novos conhecimentos.

Um outro autor, Adriano Freire (2000), menciona que todas as inovações resultam da conjugação de três requisitos básicos: oportunidade, vontade e capacidade. A oportunidade é o factor que está directamente relacionado com o mercado; a vontade relaciona-se quer com a intencionalidade na detecção de oportunidade, quer com a determinação e empenho colocados no desenvolvimento da inovação; e a capacidade, traduz-se em todo um conjunto de recursos necessários, financeiros, humanos e técnicos.

⁴ A pesquisa de informação na *WWW* será objecto de algum desenvolvimento a apresentar no Capítulo III.

Um outro aspecto a salientar, é a participação da informação no modo de fazer progredir o processo de inovação (Alberto Carneiro, 1995). É vital que a empresa constitua, mantenha e desenvolva um sistema de informação a fim de estar actualizada sobre a evolução científico-tecnológica e o lançamento de inovações. A informação assume um valor e um papel crucial no âmbito da gestão empresarial. O gestor tem de estar muito bem informado, pois a qualidade de uma decisão não pode ser melhor do que a da informação que lhe está subjacente.

Na perspectiva da *InnovationNetwork* (2003), para que a inovação nasça é necessário ter acesso a nova informação; ter uma visão clara e simples dos clientes; ter tempo para pensar e falar sobre as diversas possibilidades; ter visão, paixão, coragem e persistência; ter uma diversidade de pensamento, experiência e habilidade; ter integridade, confiança e honestidade; ter estrutura e disciplina. Estes autores consideram que as fontes de inovação se baseiam essencialmente em grandes desafios. A inovação alimenta-se da curiosidade; da criatividade; da partilha de informação; da colaboração; da liberdade para explorar; e de um fluxo interminável de ideias.

I.2.2 Tipologia da Inovação

Da análise do conjunto de ideias e definições referidas anteriormente, é possível extrair um conjunto de elementos que surgem inevitavelmente associados ao termo "*inovação*". No Livro Verde Sobre Inovação, publicado em 1996, a Comissão Europeia apresentou um conceito abrangente de inovação que enquadra todas as ideias e elementos anteriormente enunciados: "Inovação é:

- ➔ A renovação e alargamento da gama de produtos e serviços e dos mercados associados;
- ➔ A criação de novos métodos de produção, de aprovisionamento e de distribuição;
- ➔ A introdução de alterações na gestão, na organização do trabalho, bem como nas qualificações dos trabalhadores."

Nesta definição é possível observar a estruturação do conceito em torno de três blocos principais: a inovação ao nível dos produtos ou serviços, ao nível dos processos e ao nível das organizações. Estes dois últimos blocos são muitas vezes esquecidos: quando se fala de inovação, esta é na maioria das vezes associada a produtos.

A inovação ao nível dos processos (Freire, 2000) é aplicada no seio da organização. A empresa progride adaptando as tecnologias e processos à sua própria realidade e melhora-os. Em geral, este tipo de inovação consiste na utilização de tecnologias inovadoras no processo, que podem ser completamente novas ou resultar de

um processo de transferência e adaptação tecnológica a partir de outras áreas (Baptista, 1999).

A inovação ao nível dos produtos e serviços é orientada para o mercado (Freire, 2000). Este tipo de inovação, como o próprio nome indica, centra-se no desenvolvimento e melhoria das funções dos produtos, e é aquele que é mais reconhecido pelo consumidor.

Como inovar não é só comprar novas máquinas e desenvolver um produto não é só tê-lo pronto no laboratório, a inovação ao nível das organizações também é necessária. A empresa sente a necessidade de implementar sistemas de gestão de criatividade e inovação e metodologias adequadas para maximizar a sua capacidade de gerar e concretizar ideias, minimizar os riscos e aumentar a probabilidade de sucesso (Oliveira, 1999).

I.2.3 Inovação como Factor de Competitividade

A competitividade é um conceito usado em estudos de vários domínios, com perspectivas diferentes, que incluem a estratégia e a gestão, a vantagem comparativa, a competitividade ao nível do preço, ou até a perspectiva histórica e sócio-cultural. No mundo empresarial, a consideração das tecnologias de informação (TI) tem duas dimensões: a estratégica, cujo objectivo é melhorar a competitividade, reduzir a incerteza e facilitar a decisão; e a organizacional, que implica uma análise prévia do funcionamento actual, da lógica interna e da organização eventual da empresa em função das possibilidades oferecidas por estas tecnologias da informação (Carneiro, 1995). A combinação das técnicas de comunicação com os equipamentos informáticos permite que a empresa resolva rapidamente todas as necessidades que dependam da informação. As TI não consistem apenas na utilização integrada de equipamentos informáticos, e devem ser encaradas e desenvolvidas pelas empresas, numa óptica de criação e uso da informação, como instrumento indispensável aos objectivos da gestão empresarial.

Segundo Freire (2000), a emergência das novas tecnologias constitui o ponto de partida para a inovação empresarial. Tecnologias emergentes em áreas tão variadas como as telecomunicações, a produção ou os transportes estão a induzir não só o lançamento de novos produtos, serviços e processos, mas também a criação de novos negócios.

John Fitzgerald Kennedy, antigo presidente dos EUA, afirmou que *"A mudança é a lei da vida. Aqueles que olham apenas para o passado ou para o presente serão*

esquecidos no futuro”. Com base nesta frase e adaptando-a às empresas, pode-se concluir que estas não podem deixar de integrar a inovação na sua estratégia de actividade da empresa. Assim, as empresas garantem a manutenção ou o reforço da sua actividade competitiva no mercado (Baptista, 1999). Porém, importa compreender que a inovação não confere, por si só, uma vantagem competitiva, constitui antes uma potencial fonte de geração de vantagens competitivas na empresa. De facto, o impacto estratégico da inovação deriva da sua influência no reforço das competências centrais da organização e nos factores críticos de sucesso do seu negócio.

I.3 Inovação em Portugal

O tema da inovação começa gradualmente a assumir um carácter estratégico em Portugal. De facto, cada vez mais as empresas estão a chegar à conclusão que o sucesso sustentado se baseia numa permanente dinâmica de inovação, assente na renovação da oferta e das práticas de gestão (Adriano Freire, 2000).

De uma forma geral, nos estados membros da União Europeia, em que Portugal se insere, existe um número considerável de medidas e políticas de apoio e fomento à inovação. Em 1996 foi lançado, pela Comissão Europeia, o Primeiro Plano de Acção para a Inovação na Europa de que resultou o *Trend Chart on Innovation in Europe*. Trata-se de uma ferramenta com o objectivo de recolha e sistematização de informação sobre inovação, incluindo os resultados obtidos nos vários estados membros. Também o Conselho Europeu, realizado em Março de 2000 em Lisboa, desafiou a União no sentido desta alicerçar, até ao final da década, a sua economia no conhecimento. A quantificação deste objectivo surgiu em 2002, em Barcelona, com a indicação de que em 2010 a União Europeia deverá gastar 3% do seu PIB em Investigação e Desenvolvimento.

O II Inquérito Comunitário às Actividades de Inovação (Conceição e Ávila, 2001) revela que a capacidade de inovar varia grandemente de empresa para empresa, e é determinada por um vasto número de factores, tanto internos como externos à empresa: dimensão da empresa, sector de actividade, localização geográfica, ano da empresa, entre outros. O principal objectivo da inovação assinalado tanto pelas empresas industriais como pelas que prestam serviços foi a melhoria da qualidade do produto/serviço. Numa outra vertente deste inquérito, procurou-se determinar o grau de importância atribuído pelas empresas a um conjunto fixo de fontes de informação que contribuíram para a introdução e execução de projectos de inovação. A Figura 2 mostra o grau de importância atribuído a cada uma das doze fontes de informação consideradas. Verifica-se que as fontes de informação consideradas como muito importantes por um

maior número de empresas, tanto industriais como de serviços, foram “outras empresas do grupo” e “fontes de informação internas à empresa”. De salientar que a fonte de informação “redes computacionais (Internet)”, foi considerada por cerca de 70% das empresas industriais como uma fonte não relevante. No caso das empresas de serviços, as fontes de informação menos utilizadas, e consideradas importantes por menos empresas são as patentes, laboratórios de Estado, universidades e redes computacionais (Internet). No entanto, a Internet foi utilizada por quase 40% das empresas, e mais de 10% consideraram-na muito importante. Este resultado parece indicar que as empresas de serviços estão mais “abertas” do que as empresas industriais no recurso às novas tecnologias da informação como fonte de informação.

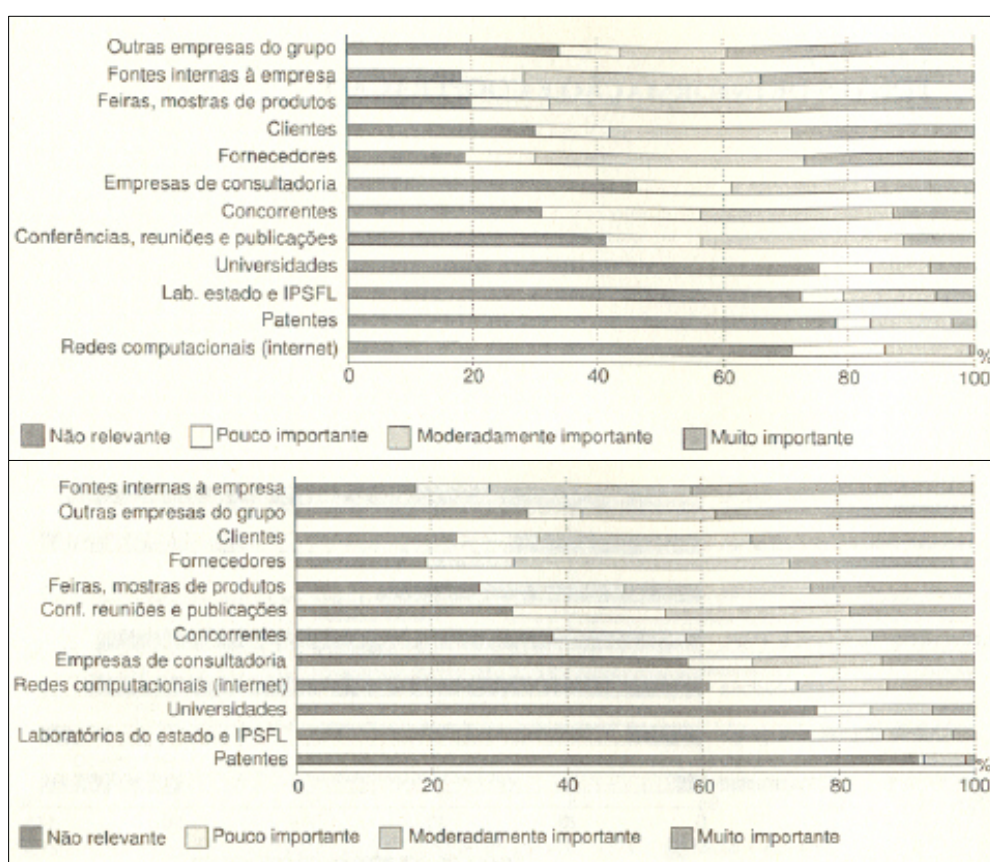


Figura 2 – Fontes de Informação para as Empresas Industriais (gráfico superior) e de Serviços (gráfico inferior). Fonte: (Conceição & Ávila 2001).

O II Inquérito Comunitário às Actividades de Inovação revela ainda uma comparação dos resultados entre Portugal e os 13 países europeus para os quais há dados comparáveis. A taxa de inovação na indústria e nos serviços em Portugal foi a mais baixa; contudo na área dos serviços, a taxa é comparável à da França e da Suécia, e superior à da Finlândia, da Noruega, e da Bélgica. Quanto às fontes de informação para a inovação, Portugal diferenciou-se na intensidade com que foram consideradas muito

relevantes pelas empresas. No entanto, as quatro fontes de informação mais importantes na Europa foram coincidentes com as quatro fontes frequentemente citadas pelas empresas portuguesas: fontes internas à empresa, clientes, outras empresas do grupo, feiras. A maior utilização e valorização das redes computacionais nos serviços, em relação ao que sucede à indústria, verificaram-se tanto para Portugal como a Europa.

Segundo Freire (2000), a generalidade das empresas portuguesas está a prestar uma crescente atenção à inovação. Um estudo deste autor (Freire, 2000, Cap.14, pág.381) sobre as práticas de gestão de inovação em pequenas e médias empresas portuguesas oriundas de sete sectores industriais permitiu reunir um conjunto restrito de factores que ajudam a promover a inovação: as características da gestão; o perfil dos empresários, o empenhamento e visão dos empresários portugueses determinam a atitude face à inovação; a abertura ao exterior; as estruturas funcionais; o estilo da liderança; o recrutamento e integração de jovens; clientes exigentes; os períodos de crise; a certificação da qualidade; e a cooperação na inovação. Na realidade, as empresas que não inovam constantemente os seus produtos, serviços e processos não conseguem gerar fortes competências nos seus negócios, acabando por nunca dispor de vantagens competitivas (Freire, 2000). Freire refere que é fundamental perceber que a batalha da inovação de processo é decisiva para inverter a tendência de decréscimo da competitividade de Portugal.

Como consequência da chamada “estratégia de Lisboa”, foi lançado em Maio de 2001 um programa com um conjunto alargado de políticas dirigidas à inovação – o PROINOV⁵. Trata-se de um programa que visa a criação de uma dinâmica de sinergias ao nível de formas sociais produtivas, concretizando-se na incorporação e desenvolvimento de práticas inovadoras. De uma forma geral pode dizer-se que os seus objectivos se estruturam em quatro dimensões:

- Promoção da inovação e da iniciativa empresarial;
- Reforço da formação e qualificação da população portuguesa;
- Impulso do enquadramento geral favorável à inovação;
- Dinamização do funcionamento do sistema de inovação em Portugal.

No que respeita a estruturas de suporte à inovação em Portugal, devem ser referidos alguns Centros Tecnológicos, como o CTC (calçado), o CENTIMFE (moldes) e o CITEV (têxtil). Também merecem destaque o Programa PME Digital e algumas acções da responsabilidade do IAPMEI e Agência de Inovação.

⁵ <http://proinov.gov.pt> (consultado em 10/09/2004).

De acordo com a Agência de Inovação, e tendo por base o *Innovation Scoreboard* 2002⁶, embora a maior parte dos indicadores de desempenho em inovação para Portugal se situe abaixo da média europeia, existem algumas áreas que apresentam resultados animadores; são exemplo a Sociedade de Informação, em particular no que respeita às despesas de TIC e ao acesso doméstico à Internet, este último incentivado por iniciativas governamentais como o POSI.

I.4 Conclusão

A inovação é qualquer coisa que transforma uma pequena faísca num fogo sustentável que, por sua vez, fortalece algo novo e com valor no mundo (*Innovation Proverbs*, 2003). Hoje em dia, uma “coisa inovadora” é algo que é muito simples de se fazer ou perceber. Mas, infelizmente, inovação não é criada com base em fórmulas mas sim por pessoas que sentiram uma necessidade de mudança.

Os fundadores de *Innovation Network* que definem inovação como: “implementação de novas ideias que gerem valor”, lançaram um desafio aos membros da Universidade de Inovação⁷, de criar um modelo de inovação. O resultado foi o *InnovationDNA* (2002) que tenta explicar o sistema ou vida da inovação. Os investigadores ao iniciarem a sua pesquisa basearam-se no DNA⁸, ou seja, na essência da vida e tentaram descobrir os elementos genéticos que formam as infinitas possibilidades da inovação (Figura 3). Nesta perspectiva, a inovação possui os seguintes elementos: o contexto, formado pela interacção entre a organização e o mundo que a rodeia; a liderança, porque a inovação pressupõe viajar pelo desconhecido o que torna necessária a confiança na visão, nas estratégias, ou nas intenções dos líderes das organizações; os valores básicos, a aprendizagem, empenho, colaboração são exemplos de princípios que definem uma organização; a cultura, que é o resultado da reflexão da liderança e dos valores da organização e que cria o ambiente necessário para agir, tomar decisões e inovar; as pessoas que constituem a fonte de inovação. Contudo os princípios fundamentais da inovação consistem em sete dimensões operacionais:

- ➔ Desafio (arranque) – inovação significa fazer algo de maneira diferente, explorando novas áreas e correr riscos;

⁶ <http://trendchart.cordis.li/scoreboards/scoreboard2002/index.cfm> (consultado em 10/09/2004).

⁷ Programa educacional desenvolvido pela *Innovation Network* que realiza “fóruns” de discussão e seminários com os seguintes objectivos: estudar as organizações que obtiveram sucesso através da prática de inovação; recolher e partilhar histórias de inovação; criar modelos que transmitam os princípios de inovação; e promover a inovação.

⁸ *Deoxyribonucleic Acid* – Esta molécula, que reproduz o código genético, é responsável pela transmissão das características hereditárias de cada espécie, quer seja nas plantas, nos animais (incluindo o homem) ou nos microrganismos.

- ➡ Clientes (impulso) – todas as inovações devem focalizar-se na criação de valor; interagir com os clientes e perceber as suas necessidades estimula o aparecimento de novas oportunidades e respectiva motivação à sua implementação;
- ➡ Criatividade (cérebro) – tudo começa a partir de uma ideia e a melhor forma para que esta seja brilhante é gerar uma infinidade de possibilidades. Enquanto a criatividade é uma característica pessoal, a capacidade de desenvolver ideias e relacioná-las com os diversos conceitos pode ser conseguida através de treino e exercício;
- ➡ Comunicação (sangue) – a circulação de informação, ideias e sentimentos transmitem vida à inovação;
- ➡ Colaboração (coração) – a inovação é um processo e alimenta-se de interacção, de informação e do poder das equipas;
- ➡ Acabamento (músculo) – as novas inovações são projectos realizados com sucesso através de processos definidos e de fortes capacidades: tomada de decisão, delegação, programação de tarefas, controlo e reacção. As organizações inovadoras sabem que a celebração também é uma parte importante de terminar um projecto, independentemente do nível de sucesso.
- ➡ Reflexão (escada) – avaliar os resultados, benefícios e custos de novos projectos é essencial para dar continuidade à inovação.

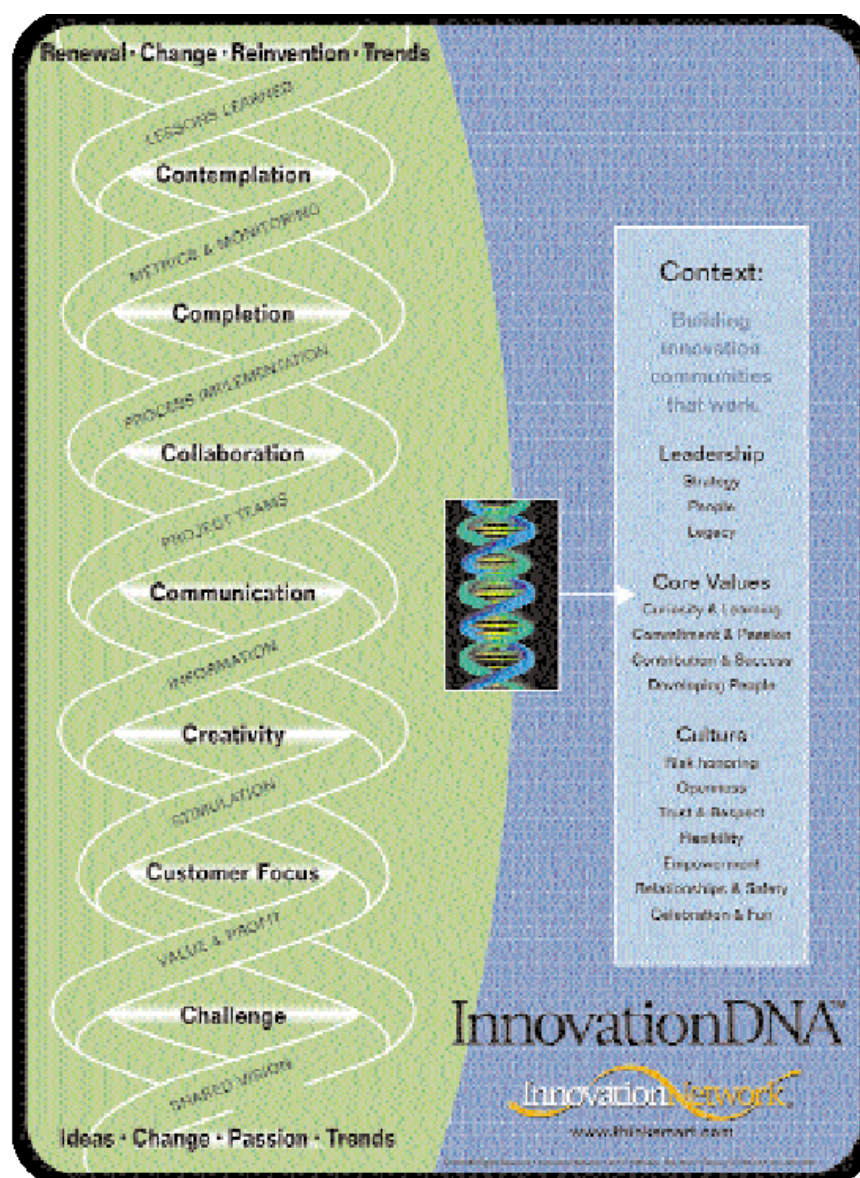


Figura 3 – Innovation DNA. Fonte: (InnovationNetwork 2002).

Da análise deste tema, pode-se concluir que nada é mais arriscado do que não inovar. O desafio presente nos dias de hoje é viver e prosperar num mundo onde a inovação é cada vez mais necessária. A criatividade tem sido uma das ferramentas de inovação cujo resultado final, quer seja um novo produto, um serviço ou um processo, tem de trazer valor para o mundo. Uma infra-estrutura de comunicação tem de existir numa organização de forma a promover e a incentivar um fluxo aberto de informação.

De facto, num tempo em que a informação concreta se desactualiza a cada segundo, é importante a sua circulação. No capítulo seguinte, vai analisar-se o modo como as pessoas pesquisam e processam informação na Internet, já que este é um

veículo que permite o acesso à informação no mesmo segundo que esta se actualiza/desactualiza.

Capítulo II – Pesquisa e Utilização de Informação

II.1 Introdução

Segundo Alberto Carneiro (1995) a informação deve constituir uma matéria-prima a partir da qual se constitua um produto útil à empresa, ao seu percurso competitivo e à sua posição pretendida no mercado. O grande desafio que se encontra e apresenta à gestão, é descobrir habilmente problemas e resolvê-los com soluções óptimas ou quase óptimas. O decisor tem de estar muito bem informado, pois a qualidade de uma decisão não pode ser melhor do que a da informação que lhe está subjacente. O gestor tem de saber seleccionar a quantidade e o tipo de conhecimentos, isto é, a informação necessária à relação com o meio ambiente. Um outro aspecto importante, que justifica a necessidade de informação devidamente recolhida e tratada, é o quanto ela pode assumir na natureza a antecipação de situações, sujeita às dimensões de uma visão probabilística. Isto é, os dados conhecidos de situações passadas podem ser inter-relacionados e darem ao gestor a consciência prognóstica de quais poderão ser os pontos principais do percurso da sua área de gestão ou da empresa.

II.2 Comportamentos Humanos na Pesquisa de Informação

Choo (2000) considera que os indivíduos vêem frequentemente a informação como algo que reside em documentos, livros, sistemas de informação e outro tipo de objectos. Choo defende que a informação não reside apenas numa forma material, mas também na mente dos indivíduos. As pessoas produzem activamente informação através dos seus pensamentos, acções e sentimentos. Como os indivíduos usam tipicamente informação para resolver um problema, executar uma tarefa, o ambiente social onde a informação é encontrada determina o seu valor e alcance. A informação pode ser tratada como um objecto ou não: como um objecto, quando estamos preocupados em adquirir informação que precisamos e como a representar em ordem a facilitar o seu acesso e processamento; quando não a tratamos como um objecto, preocupamo-nos em perceber o seu alcance comportamental e social. O facto de entender melhor a informação e a sua repercussão social, permite projectar melhor o processo da informação e os sistemas de informação. Para este autor a actividade de pesquisa de informação divide-se em três momentos: o sentimento de necessidade de informação, a pesquisa de informação e o

uso da informação. Cada um destes momentos sofre influências de factores cognitivos, afectivos e situacionais.

II.2.1 Necessidades de Informação

Estudos sobre as necessidades de informação têm vindo a ser adicionados à percepção das pessoas de como se pesquisa informação. Uma grande abundância de dados tem vindo a ser recolhidos e analisados sobre os canais, métodos e fontes usadas por grupos de cientistas, técnicos, profissionais, oficiais do governo, cidadãos e outros indivíduos que pesquisam informação. No sentido de promover a troca de informação e de simplificar o acesso à informação, várias inovações foram desenvolvidas, como por exemplo, a publicação de artigos, a personalização de ferramentas que ajudem os utilizadores, entre outros. Os estudos partilham a mesma opinião de que a informação é construída com base nos pensamentos e sentimentos das pessoas, e que esta é desenvolvida nas condições do trabalho e da vida que determinam o seu uso e o seu valor.

Brenda Dervin (1992) tem-se dedicado ao estudo de como os humanos percebem as necessidades de informação como lacunas cognitivas. Numa aproximação ao “fazer sentido” (*sensemaking*) (cfr. Choo, 2000), uma pessoa move-se pelo espaço e pelo tempo adquirindo experiências. Considerando que as suas experiências fazem sentido, ela poderá prever os próximos passos. De vez em quando o movimento é bloqueado pela percepção de uma lacuna cognitiva – uma situação que não faz sentido para a pessoa. Para ultrapassar esta lacuna cognitiva a pessoa pesquisa informação que possa resolver a situação de um modo satisfatório (que faça sentido), e usando-a possa continuar o seu trabalho. Dervin e os seus colaboradores completaram mais de 40 estudos nas duas últimas décadas baseados na aproximação ao fazer sentido (*sensemaking approach*). A sua investigação sugere que o modo como as pessoas percebem as suas lacunas cognitivas e os modos como elas procuram a informação de que necessitam são bons indicadores do seu comportamento na pesquisa de informação. Melhor ainda, as maneiras com que as pessoas percebem as suas lacunas cognitivas podem ser codificadas em categorias universais e estas serem aplicadas aos diferentes grupos de utilizadores de informação. Por exemplo, um conjunto de categorias, chamadas de *situation stops*, identificado por Dervin (1992) inclui o seguinte:

- ➔ *Decision Stop* –Where the human sees two or more roads ahead.
- ➔ *Barrier stop* – Where the human sees one road ahead but something or someone stands on the road blocking the way.

- ➔ *Spin-out stop* - Where the human sees self as having no road.
- ➔ *Wash-out stop* - Where the human sees self as on road that suddenly disappears.
- ➔ *Problematic stop* – Where the human sees self as being dragged down a road not of his or her own choosing.
- ➔ *Other situation categories* that depend on how the human judges perceptual embeddedness (how foggy is the road), situational embeddedness (how many intersections are on the road), and social embeddedness (how many people are also travelling).

As necessidades de informação são geralmente pensadas em termos das necessidades cognitivas das pessoas – lacunas ou deficiências no estado mental do conhecimento que podem ser representadas por questões colocadas ao sistema de informação ou fonte. Como a informação é procurada e usada num ambiente social, a informação tem de satisfazer para além das necessidades cognitivas, necessidades afectivas e emocionais (Tom Wilson, 1994). Quando a informação não faz sentido, a sua incompreensão gera um estado de incerteza. Carol Kuhlthau (1993) considera que a incerteza provoca vários sintomas afectivos, tais como, ansiedade, apreensão, confusão, frustração e falta de confiança. Estes estados afectivos impulsionam e direccionam a procura e a experiência na utilização da informação pela pessoa. As respostas afectivas influenciam e são influenciadas pela habilidade do indivíduo em criar conhecimento, focar as necessidades de informação, gerir expectativas e aprofundar o interesse pessoal na procura. Tom Wilson (1997) sugere que a incerteza e os seus sintomas afectivos constituem um estado de tensão (*stress*) com o qual o indivíduo tem de lidar.

Ao nível situacional, Robert Taylor (1991) identificou elementos do ambiente do uso da informação que poderiam modificar as formas estruturais e conscientes da necessidade da informação. Por exemplo, em que grau é que a situação é nova ou familiar, simples ou complexa; e a dimensão com que os participantes concordam ou não com as suposições, objectivos, e opções, que influenciam de forma idêntica a natureza e intensidade da necessidade da informação. Como resultado, para determinar as necessidades de informação não é suficiente colocar apenas a questão “O que é que tu precisas de saber?” mas também colocar questões como “Porque é que precisas da informação?”.

Susan MacMullin e Robert Taylor (1984) identificam onze dimensões que definem a necessidade de informação e servem como critério pelo qual a utilidade e o valor da informação são avaliados. Estas dimensões fornecem uma representação detalhada do

ambiente que rodeia as situações problema, e sugerem formas para definir as necessidades de informação. Na Figura 4 apresentam-se essas dimensões.

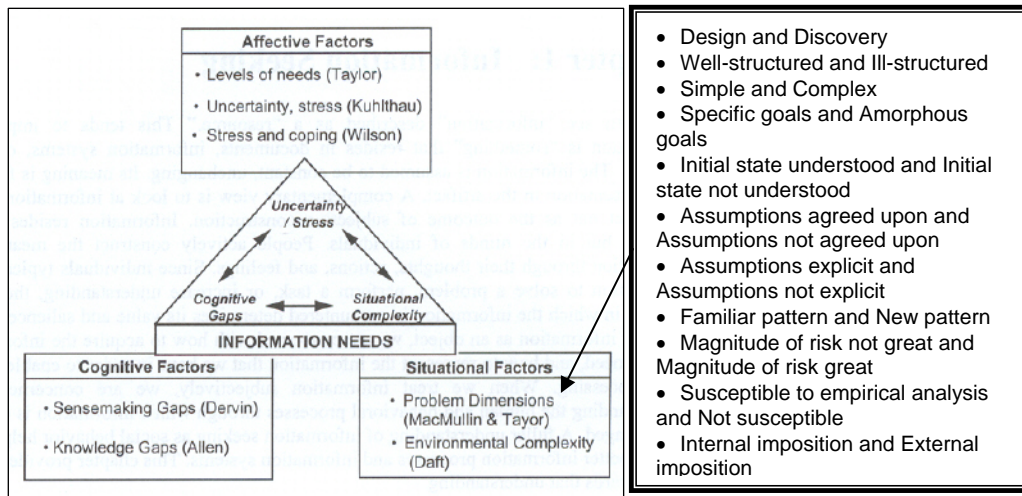


Figura 4 – Factores que afectam as necessidades de informação. Fonte: (Choo 2000).

II.2.2. Pesquisa de Informação

A pesquisa de informação é um processo no qual os humanos determinam a finalidade da pesquisa de forma a mudar o seu estado de conhecimento (Marchionini 1995). Para uma mesma necessidade de informação, cada indivíduo procura-a de uma maneira diferente, dependendo do seu conhecimento sobre fontes, experiências passadas, preferências pessoais, entre outros. Kuhlthau divide o processo de pesquisa de informação em seis etapas: iniciação, selecção, análise, formulação, recolha, e apresentação. O indivíduo identifica possíveis fontes, diferencia e escolhe algumas, localiza-as e entra em contacto com elas, interagindo de forma a obter a informação desejada (Figura 5).

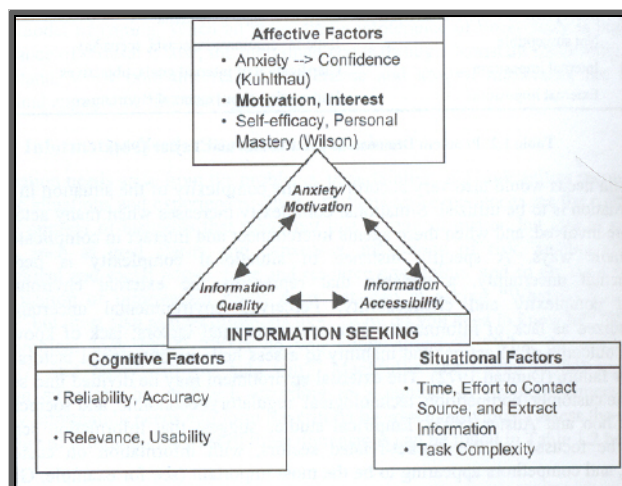


Figura 5 – Factores que afectam a pesquisa de informação. Fonte: (Choo 2000).

Hoje em dia, num ambiente rico de informação, é importante saber como é que o indivíduo selecciona as fontes e a informação proveniente de várias fontes. Uma forma de contornar esta questão é de avaliar a quantidade de esforço necessário para usar a fonte relacionando-a com a utilidade da informação fornecida por essa mesma fonte. Ao mesmo tempo, a avaliação de custos e benefícios é modelada pelo interesse pessoal e pela sua motivação, e pela dificuldade da tarefa ou do problema a resolver.

A nível cognitivo, o indivíduo selecciona uma fonte que lhe pareça ter uma maior probabilidade de fornecer informação relevante e útil. A relevância e a utilidade podem depender dos atributos da informação tais como, a fiabilidade, a utilidade e a pertinência para o problema em causa. Assim, o indivíduo irá preocupar-se com o rigor e a fiabilidade da fonte. A selecção da fonte e da informação seriam conduzidas pela percepção cognitiva e pela exactidão. As pesquisas de informação agrupam alguns ou todos estes atributos na ideia “percepção da qualidade da fonte” de forma a examinar até que ponto influencia o uso dessa fonte (Figura 6).

A nível afectivo, o grau de motivação do indivíduo e interesse pelo problema determinam a quantidade de energia que este vai investir na pesquisa da informação. Kuhlthau (1993) sugere que à medida que a pesquisa de informação vai evoluindo, os sentimentos iniciais de incerteza e a ansiedade diminuem e a confiança aumenta. Se a informação encontrada permitir ao indivíduo ver o problema de uma forma mais clara, os sentimentos de optimismo e confiança aumentarão à medida que a pesquisa evolui. Wilson (1997) defende que desde que exista um sentimento forte de auto-eficácia ou de mestria pessoal relativamente ao uso de uma fonte, este irá conduzir a um aproveitamento e desenvolvimento de competências para o uso correcto dessa fonte, cuja percepção conduzirá a uma melhor decisão em relação à pertinência do uso da informação contida nesse recurso.

A nível situacional, a selecção e uso de fontes são influenciados pela quantidade de tempo e esforço que são exigidos para localizar, contactar e interagir com a fonte no sentido de extrair informação. Os atributos da fonte podem ser integrados numa única variável chamada “percepção da acessibilidade da fonte”. Mary Culnan (1985) define esta variável como “o nível de esforço necessário para usar uma fonte particular de informação”. Esta autora propõe três dimensões de acesso à informação: a dimensão física, que consiste no acesso à fonte de informação (por exemplo, ir à biblioteca); a dimensão *interface*, que consiste em transformar a informação numa linguagem perceptível; e a dimensão informação, que consiste na extracção da informação relevante. Uma quarta dimensão poderia ser considerada, a psicológica: quando se tem

de lidar com uma fonte desagradável ou o simples embaraço de revelar a ignorância e/ou a necessidade de ajuda.

A Figura 6 resume o processo de pesquisa de informação: a selecção e o uso das fontes de informação são influenciados por dois atributos: a percepção da acessibilidade da fonte e a percepção da qualidade da fonte. Ao mesmo tempo, a complexidade da tarefa a cumprir, a incerteza do ambiente que rodeia o problema ou a motivação pessoal também influenciam a pesquisa de informação.

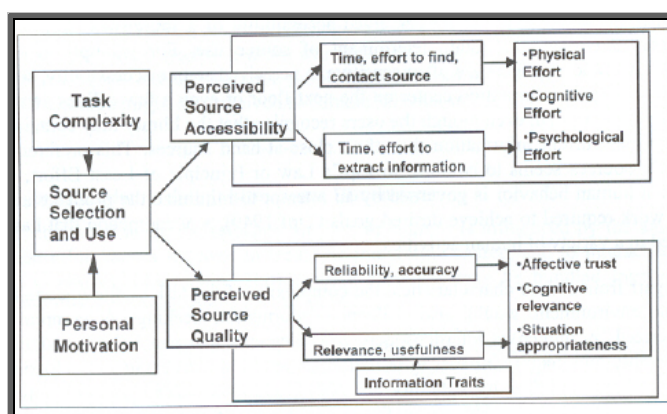


Figura 6 – Pesquisa de Informação: Acesso e Qualidade. Fonte: (Choo 2001).

II.2.3 Uso da Informação

De uma certa forma, pode dizer-se que a necessidade de informação e o uso dessa informação são os dois lados de uma moeda, desde que a informação necessária seja usada. Robert Taylor (1991) propôs oito categorias que definem o processo e uso de informação, geradas pela percepção das necessidades de informação em situações particulares. Assim, segundo este autor, a informação pode ser usada para desenvolver um assunto; proporcionar uma melhor compreensão do problema; saber o que fazer e como fazer; determinar factos para descrever a realidade; comprovar outra informação; prever acontecimentos futuros; iniciar ou motivar o envolvimento pessoal; e, por último, estabelecer relações (Figura 7).

A nível cognitivo, o estilo e preferências dos indivíduos influenciam a forma com que a informação é processada e usada. Diversos métodos e instrumentos foram desenvolvidos para diferenciar tipos de personalidade e preferências cognitivas. O *Myers-Briggs Type Indicator* (MBTI) é o instrumento mais frequentemente utilizado por

classificar os tipos de personalidade em quatro pares de características⁹. Estes quatro pares de atributos ao serem combinados criam 16 tipos de personalidade que definem, cada um deles, o processo e uso de informação de uma maneira distinta. Outras variáveis cognitivas possíveis de considerar são a dependência da área em que o indivíduo se insere e a estilos de adaptação à inovação. Os indivíduos dependentes tendem a responder, sem uma análise crítica, às sugestões da área, enquanto que os independentes se orientam correctamente sem influência do ambiente que os rodeia. Quando os indivíduos inovadores seleccionam e usam informação, estes exploram novos elementos e geram novos modelos, enquanto que aqueles que se adaptam ao sistema, apoiam os quadros de referência e tentam melhorar o sistema existente.

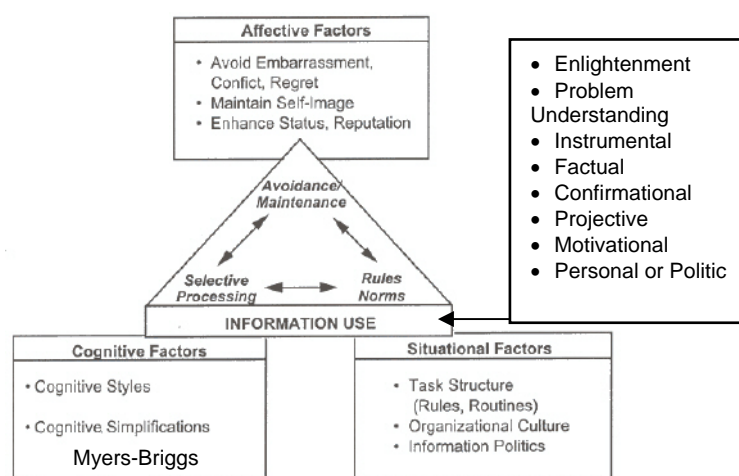


Figura 7 – Uso de Informação. Fonte: (Choo 2001).

A nível afectivo, as pessoas tendem a evitar a utilização de informação que provoque emoções fortes e negativas nos outros ou em si próprios. As pessoas adquirem e usam informação de uma forma selectiva para evitar embaraços, conflitos ou arrependimentos, para manter a auto-imagem e desenvolver a reputação pessoal (Figura 7). Por exemplo, os gestores avaliam positivamente a continuação de um projecto a longo prazo, mesmo quando a informação disponível indica que é necessário parar para reduzir

⁹ Os quatro pares de características do *Myers-Briggs Type Indicator* (MBTI) são:

- introversão *versus* extroversão: os indivíduos introvertidos obtêm a energia mental de si próprios, enquanto que os extrovertidos vão buscá-la aos outros;
- sensação *versus* intuição: uns indivíduos confiam na informação recolhida através dos cinco sentidos e outros, intuitivos, confiam mais em padrões, relações e impulsos;
- pensamento *versus* sentimento: os indivíduos pensadores tomam decisões baseadas em critérios objectivos, enquanto que os sentimentais baseiam-se em valores pessoais para decidir entre o que está certo ou errado;
- julgamento *versus* compreensão: os juizes retiram conclusões a partir da informação disponível, enquanto que os conhecedores mantêm em aberto as suas opções, retirando parte do seu tempo para juntar informação disponível.

custos adicionais. Um aspecto psicológico determinante para o uso de informação é o desejo de “salvar a pele”. Os gestores insistem porque se recusam a admitir os seus próprios erros, muito menos expô-los. Um outro exemplo é o síndrome do “*not-invented-now*”, que consiste na tendência para rejeitar informação nova vinda do exterior. Este comportamento pode ser consequência natural dos indivíduos que desenvolveram uma certa estabilidade profissional de forma a reduzir a incerteza e o *stress* que enfrentam diariamente. Como resultado, os gestores reagem emocionalmente e resistem a novas ideias e informações externas que possam fragilizar as suas decisões e convicções.

A nível situacional, o processamento e uso de informação são determinados pela influência das regras e normas do grupo social, profissional ou organizacional (Figura 7). Por exemplo, os registos e relatórios policiais definem qual informação é documentada e arquivada; as regras de abordagem à informação definem como esta deve ser filtrada; e regras de planeamento definem como a informação é usada para decidir sobre a origem da fonte. Os grupos fortemente unidos são susceptíveis ao «*groupthink*». Isto acontece quando os seus membros chegam a acordo, numa amplitude em que se comprometem relativamente ao processamento e uso de informação, e optam por ignorar a informação que intimide as convicções do grupo. As características do «*groupthink*» podem ser divididas em três grupos: os membros partilham o sentimento de invulnerabilidade que conduz a uma maior disposição para correr riscos; os membros não aceitam informações externas, pensam racionalmente e colectivamente ou ignoram informação anormal; os membros incentivam a uniformidade, ou seja, apoiam a unanimidade através da autocensura, e pressão directa sobre aqueles que discordem. Assim sendo, a cultura organizacional pode ser definida como um conjunto de suposições partilhadas, desenvolvidas pela organização, para aprender a lidar com problemas de adaptação externa e integração interna. Como resultado, a cultura organizacional desenvolve um quadro partilhado pelos membros, para que, de uma forma colectiva, a informação faça sentido. Um aspecto importante da cultura organizacional é a política organizacional. Na procura de poder, a informação pode ser usada para proteger ou expandir áreas de influência.

II.2.4 Um Modelo Integrado de Pesquisa de Informação

Os ciclos de pesquisa e uso de informação são integrados num ambiente de processamento de informação constituído por estruturas cognitivas e disposições afectivas dos indivíduos, e pelas condições sociais e profissionais. O modelo (Figura 8) destaca três importantes propriedades de pesquisa e uso de informação. Primeiro, a

forma como a informação ganha sentido depende de estruturas cognitivas e afectivas dos indivíduos. A nível cognitivo, o indivíduo estrutura o problema estabelecendo fronteiras, objectivos, relações de forma a delinear o espaço de pesquisa e uso de informação. A nível afectivo, as emoções alertam o indivíduo a assinalar algumas sugestões como sendo especialmente importantes ou a seleccionar certas fontes baseado em sentimentos que resultaram da sua experiência. Segundo, o uso de informação é situacional. Os indivíduos ao pertencerem a grupos de trabalho ou sociais, a estrutura dos problemas que o grupo costuma enfrentar, as condições em que costumam trabalhar, e o modo de definir os problemas permitem estabelecer um contexto de uso de informação. Este contexto define normas, regras e convenções que constituem comportamentos de uso de informação. Terceiro, o uso de informação é dinâmico: o processo que envolve as três etapas, a necessidade, pesquisa e uso de informação, prossegue de uma forma repetitiva. Além disso, o uso de informação também é dinâmico porque interage com os elementos cognitivos, afectivos e situacionais do ambiente. Estes elementos animam continuamente o processo de pesquisa de informação ao alterar a percepção individual da informação apropriada e ao modificar os critérios pelos quais a informação é avaliada.

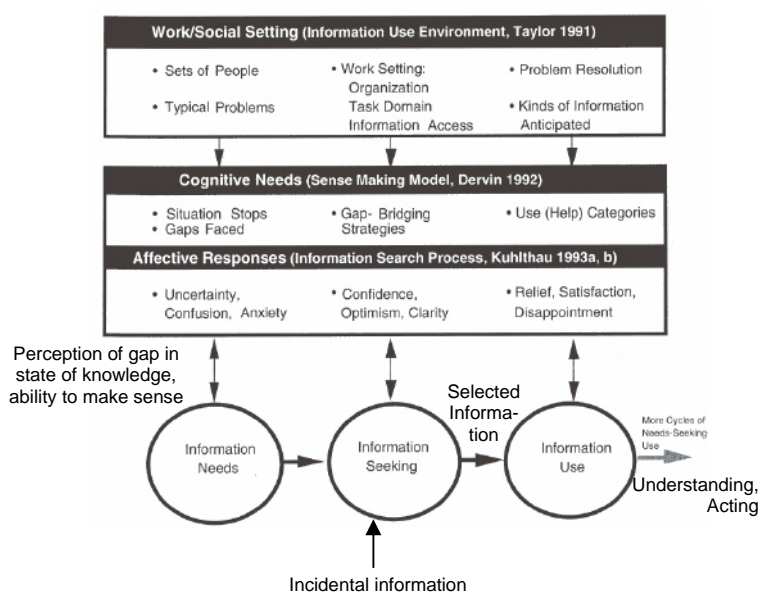


Figura 8 – Modelo integrado de uso de informação. Fonte: (Choo 1998).

Como a Figura 8 evidencia, as necessidades de informação podem ser analisadas em termos de elementos cognitivos, afectivos e situacionais. A necessidade de informação passa por vários níveis da consciência individual, desde o mais profundo do consciente até ao mais formal. A necessidade é progressivamente sentida através de

conversas com outros, observação, e reflexão até que o indivíduo seja capaz de admitir a necessidade de informação. A pesquisa de informação é a próxima etapa do modelo que se torna importante depois de o indivíduo ter desenvolvido uma compreensão clara da necessidade de informação. Características sociais ou profissionais do indivíduo podem induzir ou forçar alguns padrões de comportamento de pesquisa de informação. A cultura e a estrutura da organização, ou do grupo de trabalho, influenciam igualmente as atitudes individuais do acto de pesquisa de informação. O uso de informação é a última etapa do modelo que acontece quando o indivíduo actua com informação recolhida durante a pesquisa para, por exemplo, responder a uma questão, resolver um problema, tomar uma decisão, negociar uma posição ou fazer com que uma dada situação tenha sentido. Assim, o uso da informação é uma mudança do estado de conhecimento do indivíduo e da sua capacidade de actuar. Como já foi referido, o uso de informação é um processo contínuo e repetitivo, e durante este, novas ideias surgem assim como novas necessidades de informação.

II.3 Conclusões

O modelo de comportamento de pesquisa de informação apresentado possui várias sugestões para incrementar o desenvolvimento de sistemas de informação e da gestão de informação das organizações, de entre as quais se salientam as seguintes:

- ➔ Os sistemas de informação não devem ter apenas como objectivo responder a questões, mas também fornecer informação útil e aproveitável que ajude os indivíduos a resolver problemas e a lidar com as exigências de problemas específicos. Os projectistas de sistemas de informação para além de necessitarem de analisar fluxos de dados, precisam de desenvolver algoritmos que extraiam informação. É necessário perceber como é que a informação é construída, como o contexto do trabalho organizacional é negociado e como o valor e utilidade da informação é determinada.
- ➔ Desenvolver o conhecimento sobre a pesquisa e uso de informação, incluindo uma compreensão dos estilos e limitações cognitivas, e as formas com que as defesas emocionais podem bloquear a aprendizagem. A diversidade cognitiva fortifica uma organização, de tal forma que estabelece uma mistura de estilos, capacidades e sensibilidades para aumentar a eficiência do processamento de informação.
- ➔ Todos os indivíduos devem aprender a gerir a quantidade e a qualidade da informação. Os projectistas precisam de perceber como os indivíduos valorizam a

informação e os utilizadores necessitam perceber como avaliam a qualidade das fontes e da informação.

- ➔ Desenvolver uma cultura organizacional que valorize e encoraje a partilha de informação. Por vezes, algumas das melhores fontes são os colegas da organização e, por isso, é necessário encorajar os membros a partilharem informação livremente. A *Web* assume aqui um papel importante na partilha de informação e do conhecimento a nível mundial.

Capítulo III – Internet e Informação

III.1 A Internet e o seu Desenvolvimento

A Internet revolucionou o computador e as comunicações mundiais de uma forma que nada até hoje os tinha influenciado dessa maneira. A invenção do telégrafo, do telefone, da rádio e do computador estabeleceram etapas para esta integração de capacidades sem igual. A Internet é, ao mesmo tempo, um meio de comunicação mundial, um mecanismo de divulgação de informação, e um meio de colaboração e de interacção entre indivíduos e os seus computadores sem preocupação da localização geográfica. A Internet representa um exemplo de sucesso de investimento constante em desenvolvimento de infra-estruturas de informação. O governo, a indústria e as universidades têm sido parceiros na evolução e desenvolvimento desta nova e estimulante tecnologia.

A Internet teve a sua origem nos Estados Unidos da América, nos anos 60 do século passado, na sequência da atribuição à agência ARPA (*Advanced Research Projects Agency*), pelo Departamento de Defesa dos Estados Unidos, da tarefa de criar uma rede de computadores capaz de pôr em comunicação centros geograficamente afastados (Silva, 1999). A rede ARPANET surgiu assim com uma utilização virada sobretudo para fins académicos e militares. Inicialmente, esta rede ligava apenas quatro centros dos Estados Unidos da América; no entanto, tornou-se a primeira rede de redes, que haveria de estar na origem da Internet. Foi no âmbito das equipas de trabalho que desenvolveram a ARPANET que surgiram os protocolos TCP/IP (*Transmission Control Protocol / Internet Protocol*)¹⁰ – a tecnologia de redes que ainda hoje está na base do funcionamento da Internet, após, é claro, inúmeros desenvolvimentos a vários níveis. Na década de 80, a ARPANET deu origem a duas outras redes: a MILNET, orientada exclusivamente para fins militares e a NSFNET (*National Science Foundation*, EUA), orientada fundamentalmente para fins científicos. Foi no âmbito da NSFNET que, durante a década de 80, começou a crescer o grande sistema de redes que viria a ser conhecido como Internet. Se o primeiro grande impulso na evolução da Internet residiu na sua abertura às universidades, a nível mundial, o segundo motor da sua expansão resultou do interesse nos seus serviços por parte de muitas instituições comerciais que viram

¹⁰ Estes dois protocolos estão na base da arquitectura da Internet. O protocolo TCP funciona apenas nos computadores da rede e controla a entrega correcta das mensagens enviadas, enquanto que o protocolo IP funciona nos computadores ligados à rede onde desempenha o papel fundamental de encaminhar as mensagens através das sub-redes que constituem a rede global.

nesse novo meio de comunicação um amplo mercado a explorar. Em meados dos anos 90, a Internet começou a ser utilizada para atingir os mais variados objectivos, tais como:

- ➔ enviar correio electrónico, utilizando o SMTP (*Simple Mail Transfer Protocol*);
- ➔ transferir ficheiros de dados ou programas de computador para computador, utilizando o FTP (*File Transfer Protocol*);
- ➔ conseguir o acesso a outros computadores que possuem informação relevante, utilizando o *Telnet Protocol* (Emulação de terminal);
- ➔ criar “fóruns” de discussão.

Como factor decisivo na popularização da Internet, em 1993, surgiu a *World Wide Web*, WWW, ou simplesmente, *Web* – um sistema de páginas de hipertexto¹¹ e multimédia à escala mundial acessível em qualquer computador por meio de interfaces gráficos. Este sistema de divulgação de informação foi desenvolvido no CERN (*Conseil Europeene de Recherche Nucléaire*), actual Laboratório Europeu de Física das Partículas em Genebra. A Web tornou a Internet uma fonte de fácil acesso e de pesquisa de informação através de *software* de “navegação” na *Web* – os *Web Browsers* (como por exemplo, o *Internet Explorer* ou o *Netscape Navigator*).

Hoje em dia, a Internet é uma infra-estrutura de informação generalizada, cujo protótipo inicial era chamado de infra-estrutura nacional (ou global) de informação (Leiner 2003). A sua história é complexa e envolve vários aspectos, tecnológicos, organizacionais e populacionais. A sua influência chega não só à tecnologia das comunicações entre computadores mas também à sociedade, na medida em que os indivíduos tiram benefício das ferramentas da Internet para realizar comércio electrónico, adquirir informação, e para comunicar entre si. Assim, a utilização da Internet, em quase todos os aspectos da vida quotidiana, quer seja para obter informação de apoio a qualquer estratégia de gestão, quer seja para efectuar uma simples reserva num hotel, representa um fenómeno que não era possível imaginar aquando do seu início. Os últimos anos têm registado um crescimento galopante da Internet como ferramenta essencial de comunicação e de difusão de informação. Segundo o Inquérito piloto à Utilização das Tecnologias da Informação e da Comunicação nas Empresas, realizado pelo Instituto Nacional de Estatística (INE) em colaboração com a Unidade de Missão

¹¹ Segundo definição de Encyclopedia.com hipertexto é “a técnica para organizar bases de dados computacionais ou documentos que facilitem a extracção de informação. Dados ou informações relacionadas são conectadas através de ligações pré-definidas ou personalizadas que permitem ao utilizador seguir caminhos através da base de dados. Os dados podem ser texto, gráficos, áudio, ou vídeo (...)”.

Inovação e Conhecimento, em 2003, 82% das empresas portuguesas com dez e mais trabalhadores utilizaram computador. A proporção de empresas com dez e mais pessoas ao serviço que dispõe de ligação à Internet foi, no mesmo ano, de 70%, enquanto que a posse de *website* ascendeu a 26% (Quadro 1).

Quadro 1 – Empresas que utilizam computador, empresas que utilizam Internet e empresas que possuem *website*, segundo o número de pessoas ao serviço (%). Fonte: (INE 2003).

Escalão de pessoal ao serviço	Utilização de computador	Utilização de Internet	Posse de <i>website</i>
Total	82	70	26
10 a 49 pessoas ao serviço	80	67	22
50 a 249 pessoas ao serviço	92	87	43
250 e mais pessoas ao serviço	100	98	69

Os níveis mais elevados de utilização de computador e de Internet e de posse de *website* são detidos pelas grandes empresas – empresas com 250 e mais trabalhadores – estes valores indicam, respectivamente, 100%, 98% e 69%. Por outro lado, as empresas cujo número de trabalhadores é superior a 9 e inferior a 50 apresentam os seguintes valores de utilização destas tecnologias: cerca de 80% utilizam computador; 67% usam Internet; 22% possuem *website*.

III.2 Modelos de Pesquisa de Informação na Web

A pesquisa de informação é uma actividade importante na sociedade humana. O acesso à informação foi evoluindo ao longo dos tempos, desde a época em que a informação era transmitida apenas oralmente até aos dias de hoje, onde as novas tecnologias têm uma forte influência. Neste capítulo pretende-se estudar a extensa utilização da *World Wide Web* e o modo como as pessoas actuam quando usam a *Web* como fonte e canal de informação. Assim, irão abordar-se modelos de pesquisa de informação que surgiram com esta influência e que se centram essencialmente no comportamento das actividades de pesquisa de informação. Existem vários modelos de pesquisa de informação: de Ellis, de Kuhlthau¹², de Belkin¹³; de Marchionini¹⁴ e o modelo

¹² Kuhlthau (1991) indica um modelo que se centra na perspectiva do utilizador e considera que o processo de pesquisa de informação se define em seis fases: 1-iniciação, caracterizada por sentimentos de incerteza; 2-selecção, caracterizada por sentimentos de optimismo que permite seleccionar as fontes; 3-exploração, caracterizada pela investigação com o objectivo de reduzir os sentimentos de incerteza; 4-formulação, junção do processo com a informação encontrada acompanhada de um aumento de confiança; 5-recolha, interacção do sistema de informação com os sentimentos de confiança; 6-apresentação, que completa o processo com um sentimento de confiança ou falha dependendo da utilidade da informação adquirida.

¹³ Belkin (1993) apresenta uma visão sobre o processo de pesquisa de informação orientada para as seguintes tarefas: 1-pesquisa de uma fonte de informação; 2-aprendizagem, aumento da compreensão do problema através da selecção de informação; 3-reconhecimento, identificação de informação relevante; 4-metainformação, interacção com os itens que limitam a tarefa.

integrador desenvolvido por Choo, Detlor & Turnbull. Destes modelos, serão aprofundados o modelo de Ellis e o modelo integrado de Choo, Detlor & Turnbull.

III.3 Modelo de Ellis de Pesquisa de Informação

Ellis (1989) propõe e elabora um modelo de comportamento de pesquisa de informação baseado em estudos sobre o método de pesquisa aplicado pelos cientistas, investigadores físicos, químicos e engenheiros. Uma versão do modelo descreve seis categorias/actividades de pesquisa de informação: iniciação, ligação, pesquisa, diferenciação, monitorização e extracção. Embora o modelo de Ellis seja baseado no estudo do comportamento de académicos e investigadores, também pode ser aplicado a outros grupos de utilizadores. Ellis afirmou que estas categorias/actividades podem ser aplicadas em sistemas de informação com ambiente de hipertexto. Considerando a *World Wide Web* como um sistema de informação em hipertexto, a maior parte das actividades descritas no modelo de Ellis já são implementadas pelas capacidades presentes em qualquer *software* de navegação na *Web*. As actividades consistem em:

- ➔ iniciação: uso de um motor de pesquisa para localizar fontes de interesse;
- ➔ ligação: ligações para outras fontes relacionadas, em ambos os sentidos (ligação para a frente e para trás);
- ➔ pesquisa: sondagem das páginas *Web* das fontes seleccionadas;
- ➔ diferenciação: marcação das fontes úteis para futura referência ou visita;
- ➔ monitorização: subscrição de serviços que alertem o utilizador de nova informação ou desenvolvimentos;
- ➔ extracção: explorar toda a informação disponível numa fonte específica ou um tópico em particular.

A iniciação abrange aquelas actividades que formam a pesquisa inicial de informação – identificar fontes de interesse que poderiam servir de ponto de partida da pesquisa. A identificação de fontes geralmente abrange aquelas que são conhecidas e que já tinham sido usadas assim como as menos conhecidas das quais se espera obter informação relevante. Durante a pesquisa das fontes iniciais, estas, normalmente, podem conduzir o utilizador a outras fontes e referências. No caso da *Web*, a actividade de iniciação pode basear-se na consulta de um motor de pesquisa, por exemplo, *Google*, para encontrar uma lista de ligações relacionada com o tema de pesquisa, nos ficheiros

¹⁴ Marchionini (1995) considera o processo de pesquisa de informação em oito fases que se desenvolvem em paralelo: 1-reconhecer e aceitar um problema de informação; 2-definir e perceber o problema; 3-escolher um sistema de procura; 4-formular uma questão; 5-executar a pesquisa; 6-analisar os resultados; 7- extrair informação relevante; 8-reflectir, repetir e/ou parar.

de perguntas frequentemente elaboradas (FAQs-*Frequently Asked Question Files*), e/ou em sítios de referência.

A ligação é uma técnica que liga a fonte inicial a outras fontes que contribuem com novos recursos de informação. Por exemplo, as técnicas mais usuais consistem em seguir referências de um determinado artigo ou uma pesquisa de literatura de outros artigos referidos no artigo inicial. Existem dois tipos de ligação: “para trás” (*backward*) e “para a frente” (*forward*):

- ↳ “para trás” acontece quando as referências da fonte inicial são consultadas, e é estabelecida uma rotina de pesquisa de informação entre cientistas e investigadores;
- ↳ “para a frente” acontece quando se identifica e se segue outras fontes que referem a fonte inicial.

As desvantagens desta “navegação” são o tempo disponível e a confiança em seguir o caminho da pesquisa. Por exemplo, ao utilizar-se um *software* de navegação da *Web*, a ligação “para trás” consistiria em seguir ligações presentes na página inicial; a ligação “para a frente” poderia envolver o uso de um motor de pesquisa para encontrar outras páginas *Web* que se liguem à página inicial.

A partir do momento em que as fontes de informação foram localizadas, a pesquisa é a actividade de procura de informação nessas potenciais fontes. Por vezes, o indivíduo simplifica a pesquisa ao olhar os diversos documentos com listas de conteúdos, listas de títulos, tópicos e nomes das organizações ou pessoas. A actividade de pesquisa consiste em estar aberto a descobertas interessantes, encontrar novas ligações e aprender, o que pode causar mudanças na necessidade de informação. Na *Web*, a pesquisa é feita de uma forma natural através de um simples “clique” no rato. Com a diversidade de ligações e uma velocidade de acesso adequada, seguir uma nova ligação é extremamente fácil. A única preocupação é perder-se no oceano de ligações que pode implicar o abandono da pesquisa na *Web*. Um bom exemplo de pesquisa na *Web* seria encontrar um artigo de um jornal *online* e seguir as suas ligações para o índice geral e para um outro artigo completamente diferente. Isto por sua vez poderia conduzir a uma página de ligação com os autores da revista, o seu quadro editorial, ou páginas de organizações.

A diferenciação consiste na selecção das fontes através da indicação das diferenças entre elas e o valor da informação que cada uma oferece. Esta actividade poderia consistir em organizar as fontes segundo o tópico, perspectiva, ou nível de detalhe. A diferenciação é altamente dependente da experiência anterior do indivíduo

com as fontes ou do conselho de contactos pessoais. No ambiente da *Web*, esta actividade consiste, por exemplo, em organizar as fontes em categorias por tópicos e depois ordená-las pela profundidade da informação que contêm.

A actividade de monitorização consiste em manter actualizados os desenvolvimentos de uma determinada área através da consulta regular de fontes específicas. Os indivíduos, ao concentrarem-se num pequeno número daquilo que podem ser consideradas fontes potenciais, podem descobrir importantes desenvolvimentos de um determinado tópico. Uma actividade de monitorização do *software* de navegação da *Web* poderia ser retornar a uma fonte assinalada para ver se a página tem sido actualizada ou visitar regularmente um sítio de um jornal ou revista quando estiver prevista a publicação da sua nova edição.

A extracção é uma actividade de análise sistemática de uma ou mais fontes particulares com objectivo de identificar informação relevante. Numa pesquisa, a extracção pode ser conseguida através da consulta directa da fonte, ou indirectamente através da observação de bibliografias, índices ou bases de dados na *Web*. No caso da pesquisa na *Web*, esta actividade pode consistir em guardar a página como um ficheiro, imprimi-la para arquivo ou para anexar a um documento.

Conforme referido anteriormente, Ellis (1989) considerou que os sistemas baseados em hipertexto teriam a capacidade de implementar funções indicadas pelo modelo comportamental. Na Tabela 2, apresenta-se uma comparação entre os passos de pesquisa na literatura com os passos de pesquisa no ambiente *Web*.

Modos de Pesquisa de Informação	Passos de Pesquisa na Literatura	Passos de Pesquisa na Web
Iniciação	Identificar fontes de interesse	Identificar sítios <i>Web</i> ou páginas que podem conter informação relevante
Ligação	Seguir referências descobertas numa determinada fonte	Seguir ligações das páginas iniciais para outros sítios <i>Web</i> com conteúdo relacionado
Pesquisa	Analisar os conteúdos de tabelas ou títulos	Analisar páginas de nível superior: listas, títulos, mapas de páginas
Diferenciação	Aceder ou limitar informação de acordo com a sua utilidade	Seleccionar páginas úteis através de marcação, impressão, cópia, etc; escolher um sítio <i>Web</i> diferenciado e pré-seleccionado
Monitorização	Receber regularmente relatórios ou sumários das fontes seleccionadas	Receber actualizações de sítios <i>Web</i> ; visitar os sítios <i>Web</i> preferidos
Extracção	Trabalhar sistematicamente uma determinada fonte para identificar informação relevante	Procurar sistematicamente um sítio <i>Web</i> para obter informação de interesse

Tabela 2 – Actividades de Pesquisa de Informação. Fonte: (Choo et al. 2000).

Em estudos posteriores, Ellis (1991) adicionou mais duas categorias ao seu modelo: verificação, onde a qualidade da informação é verificada e a finalização, que consiste na conclusão do processo de pesquisa de informação, tal como a construção do sumário final e organização de notas. Estas mudanças não só reflectem estudos posteriores, mas também tornaram o processo de pesquisa de informação mais mecânico. Apesar de tornar os processos de pesquisa e as relações entre as características do modelo mais refinados, Ellis também concordou que as ligações entre as categorias se tornaram mais curtas. No caso do uso da *Web*, a actividade de verificação pode envolver a extracção de palavras-chave de uma fonte e a procura de informação que a comprove numa outra página *Web*. Contudo, a grande percentagem de informação não referenciada presente na *Web*, torna a verificação da informação muito difícil. Por vezes, a informação é verificada recorrendo a fontes tradicionais e não, propriamente, a outras páginas *Web*.

Em 1997, Ellis modificou o seu modelo, alterando a categoria iniciação para estudo. Esta categoria vem dar mais ênfase à obtenção de uma visão mais alargada da pesquisa ou a localização de pessoas chave que trabalhem nessa área. A diferenciação foi alterada para distinção, onde as fontes de informação são ordenadas. A actividade distinção inclui também a fonte de onde a informação vem. Ellis refere que as fontes informais, como as discussões ou conversas, são normalmente consideradas de maior

valor que artigos completos, bem como as fontes secundárias, como é o caso das tabelas de conteúdos e sumários. Isto acontece pelo facto do uso das fontes electrónicas ter aumentado e pela grande quantidade de informação que possui. Para o utilizador é importante que seja formada uma lista hierarquizada de resultados no processo de pesquisa de informação. Com a *Web*, é tão fácil descobrir uma fonte de informação (um *site Web* de uma organização) como difícil de a confirmar (uma fonte incluída numa página pessoal da *Web*) devido à facilidade de alteração e apresentação de informação na *Web*.

Uma outra categoria foi acrescentada ao modelo de Ellis – filtragem, que se baseia em critérios pessoais ou mecanismos que aumentem a precisão e relevância da informação. A pesquisa por restrição a um tema ou a uma palavra-chave são exemplos típicos desta categoria. Esta ideia de incluir a filtragem no seu modelo, surgiu pelo facto de que é um elemento crucial no processo de pesquisa de informação. Numa pesquisa *Web*, a filtragem consistiria na restrição da procura à data de publicação ou à anotação do URL¹⁵ da página (usando um motor de pesquisa na *Web* ou um determinado *site*). Quando combinada com a distinção, onde as fontes estão ordenadas e classificadas, começa-se a constatar como a recuperação de informação é indirectamente mencionada no modelo de pesquisa de informação de Ellis. A Figura 9 ilustra o modelo de pesquisa de informação de Ellis. É de salientar que a estrutura global do processo de pesquisa de informação poderia estar contida dentro de cada actividade, devido à natureza fractal do processo.

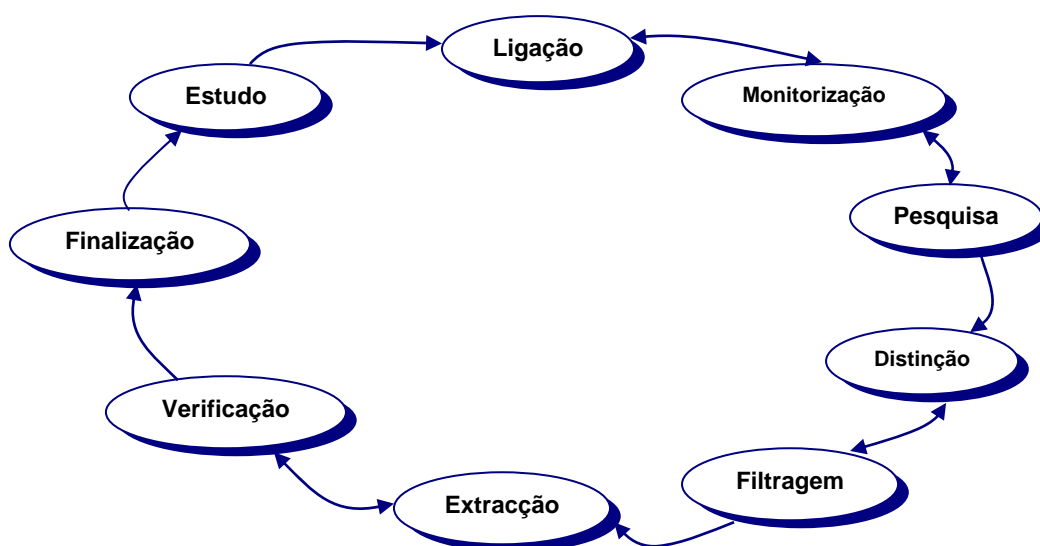


Figura 9 – Modelo de Pesquisa de Informação de Ellis.

¹⁵ URL – *Uniform Resource Locator*.

O processo de pesquisa de informação é, por natureza, como um fractal. Por exemplo, dentro da actividade de estudo, deve com toda a certeza existir ligação, pesquisa, diferenciação, e a finalização que formaliza a conclusão da actividade. Como acontece com um fractal, a mais pequena mudança numa sub-característica pode causar impacto não apenas numa característica próxima mas, também, em todo o processo de pesquisa de informação. Isto é mais do que refinar a pesquisa; as várias características da pesquisa de informação podem assumir uma apresentação diferente segundo as mudanças do pesquisador, das fontes ou da tecnologia. Estas alterações colaboram nas três actividades (ligação, pesquisa e diferenciação) onde a pesquisa de informação pode ser melhorada. Por exemplo, a colaboração entre pesquisadores é a área mais evidente de melhoria do processo. Diferentes utilizadores podem partilhar descobertas anteriores ou colaborar entre si para minimizar o trabalho futuro. À medida que as fontes ficam disponíveis digitalmente, torna-se mais fácil ligá-las e partilhá-las. A evolução da tecnologia pode permitir uma monitorização automática; combinação e comparação de resultados; e distribuição de perfis de utilizador ou programas que possam fornecer pontos de partida para a pesquisa de informação.

Ironicamente, a tecnologia torna os recursos abundantes mas também se tornam mais difíceis de serem classificados. As exigências da fonte de publicar informação são muito menores que a classificação directa e especializada, e muitas vezes exclui a indexação. Sem uma organização comum entre as fontes electrónicas, o trabalho individual será necessário para se construir um planeamento de pesquisa de uma área específica. Neste caso, a filtragem pode ajudar a estabelecer classificações operacionais de informação de comunidades de utilizadores.

III.4 Modelo de Comportamento de Pesquisa de Informação na Web (Choo, Detlor & Turnbull, 2000)

Choo e Auster (1993) fizeram uma crítica aos estudos existentes que trouxe um novo ar a todo o ambiente do utilizador. Estes autores não estudaram apenas o sistema de informação, mas também o compararam com as outras fontes de informação. Os quatro métodos de Aguilar (1967) fornecem um esquema do ambiente ao definirem quatro modos de procura:

- I. Visionamento Indirecto – o indivíduo é exposto à informação sem uma necessidade específica em mente. O objectivo de visionamento é precisamente mostrar desenvolvimentos significativos ou itens que gerem novas necessidades de informação. Tácticas típicas deste modo envolveriam a observação de várias fontes,

tirariam vantagem daquilo que é de fácil acesso, e incluiriam fontes que, no início, poderiam não parecer directamente relacionadas com o tema de pesquisa;

- II. Visionamento Condicionado – o indivíduo é exposto à informação directamente através da selecção de tópicos ou de certos tipos de informação. O objectivo principal é avaliar a relevância da informação encontrada de forma a analisar o impacto natural na organização. O observador é sensível à informação sobre estes tópicos, e é capaz de analisar, de uma forma geral, a veracidade da informação encontrada. Para aumentar o conhecimento destes tópicos, as tácticas usuais poderiam envolver pesquisa nas fontes que o utilizador conhece e que têm uma forte possibilidade de conterem informação relevante;
- III. Pesquisa Informal – o indivíduo pesquisa activamente para aumentar o seu conhecimento e compreensão de um determinado tópico. Diz-se informal porque o indivíduo não deseja despende muito tempo e esforço. O objectivo principal consiste em aprender mais sobre um tema de forma a satisfazer a necessidade de resposta;
- IV. Pesquisa Formal – o indivíduo faz um esforço deliberado e planeado para obter uma informação específica ou tipos de informação sobre um tópico em particular. A pesquisa é formal porque é estruturada segundo procedimentos e métodos pré-estabelecidos. O objectivo principal é encontrar informação relevante sobre um tópico, de forma a fornecer uma base de desenvolvimento para atingir a decisão.

Os quatro modos de pesquisa são comparados segundo as seguintes categorias: necessidade de informação, pesquisa de informação, uso de informação, esforço e quantidade de fontes, na Tabela 3.

Características	Modos de Pesquisa de Informação			
	Visionamento Indirecto	Visionamento Condicionado	Pesquisa Informal	Pesquisa Formal
Necessidade de Informação	Áreas gerais de interesse; necessidade específica	Capaz de reconhecer tópicos de interesse	Capaz de formular simples questões	Capaz de especificar objectivos em detalhe
Pesquisa da Informação	“ <i>Limpeza</i> ” – analisar a diversidade de fontes, tirando vantagem das que são de fácil acesso	“ <i>Distinção</i> ” – pesquisar as fontes pré-seleccionadas ou tópicos de interesse pré-especificados.	“ <i>Satisfação</i> ” – pesquisa centra-se apenas numa área ou tópico, mas uma pesquisa suficiente é satisfatória.	“ <i>Optimização</i> ” – Pesquisa sistemática de informação sobre um tópico, seguindo um método ou procedimento
Uso da Informação	“ <i>Pesquisar</i> ” – descoberta accidental	“ <i>Aprender</i> ” – Aumentar o conhecimento sobre um tópico de interesse	“ <i>Seleccionar</i> ” – Aumentar o conhecimento numa área especializada	“ <i>Extracção</i> ” – uso formal de informação para a tomada de decisão
Esforço/Quantidade de Fontes	Mínimo esforço e máximo de fontes possíveis	Pequeno esforço e poucas fontes	Médio esforço e poucas fontes	Grande esforço e muitas fontes

Tabela 3 – Quatro modos de pesquisa de informação. Fonte: (Choo, Detlor & Turnbull, 2000).

Aguilar alterou o centro da sua atenção para as fontes de informação e a forma como são usadas. Os quatro modos de pesquisa de Aguilar salientam que num ambiente real, a escassez de fontes inibe a continuação da pesquisa com procuras formais. A diversidade de abordagens mais eficientes exige, como prioridade, necessidades contínuas de mudança de informação. Este ponto de vista permite moldar os sistemas de informação de forma a trabalhar em harmonia com o utilizador considerando a visão total de recursos de informação no ambiente. O olhar-se não apenas para a informação encontrada no sistema mas também para a informação partilhada por outros utilizadores, poderá ajudar a avaliar a quantidade de colaboração existente na pesquisa de informação.

Choo, Detlor & Turnbull (2000) consideram que os modos de pesquisa de Aguilar e as categorias do modelo de pesquisa de Ellis podem ser combinadas e estendidas a um novo modelo de comportamento de pesquisa de informação na *Web*. A Tabela 4 identifica quatro modos de pesquisa de informação na *Web*: visionamento indirecto, visionamento condicionado, pesquisa informal e pesquisa formal, e indica quais as actividades de pesquisa de informação que podem ocorrer em cada um dos modos de pesquisa.

Actividades	Modos de Pesquisa de Informação			
	Visionamento Indirecto	Visionamento Condicionado	Pesquisa Informal	Pesquisa Formal
Iniciação	Identificar e seleccionar <i>sites</i> iniciais ou pré-seleccionados	x	x	x
Ligação	Seguir ligações presentes nas páginas iniciais	x	x	x
Pesquisa	x	Pesquisar páginas de entrada, títulos e mapas de <i>sites</i>	x	x
Diferenciação	x	Marcar, imprimir, copiar; ir directamente ao <i>site</i> conhecido	Marcar, imprimir, copiar; ir directamente ao <i>site</i> conhecido	x
Monitorização	x	Revisitar os <i>sites</i> marcados ou preferidos em busca de nova informação	Revisitar os <i>sites</i> marcados ou preferidos em busca de nova informação	Revisitar os <i>sites</i> marcados ou preferidos em busca de nova informação
Extracção	x	x	Usar motores de pesquisa (local) para extrair informação	Usar motores de pesquisa para extrair informação

Tabela 4 – Modos e Actividades de Pesquisa de Informação na Web.

No caso do visionamento indirecto na Web, existem vários exemplos da aplicação das actividades iniciação e ligação. A iniciação ocorre quando os utilizadores começam a pesquisar na Web com base em páginas seleccionadas anteriormente, ou quando visitam uma página favorita para iniciar o visionamento (como por exemplo, jornais, revistas). A ligação ocorre quando os utilizadores reparam em itens de interesse, e seguem as ligações de hipertexto para aceder a mais informação sobre esses itens. Ligação “para a frente”, do tipo já descrito, é a mais usada no decorrer do visionamento indirecto. A ligação “para trás” também é possível, desde que os motores de pesquisa possam ser utilizados para localizar outras páginas Web que apontem ao *site* em que o utilizador se encontra.

No modo de visionamento condicionado da Web verifica-se que a pesquisa, a diferenciação e a monitorização são actividades comuns na procura de informação. A diferenciação ocorre quando o utilizador selecciona páginas Web das quais espera obter informação relevante. As páginas Web podem ser diferenciadas com base em critérios

pessoais como visitas anteriores ou recomendação dadas por outros. Normalmente, os *sites* diferenciados são marcados e ao serem visitados os utilizadores pesquisam o conteúdo consultando os índices ou esquemas de *sites*. Os utilizadores podem também avaliar regularmente *sites* altamente diferenciados através da consulta repetitiva de páginas de referência ou tentar manter-se actualizados sobre a área em que trabalha (por exemplo, subscrição de *newsletters* que fornecem informação sobre novas actualizações).

Durante a pesquisa informal na *Web* é natural que as actividades de diferenciação, monitorização e extracção ocorram. A pesquisa informal é mais tentada num pequeno número de *sites* da *Web* que foram diferenciados pelo utilizador, baseado no seu conhecimento individual destes *sites* sobre a relevância, qualidade, ligação e dependência da informação. A extracção é relativamente “informal” no sentido de que a pesquisa seria apenas no(s) *site(s)* seleccionado(s). A extracção também pode fazer uso dos comandos de pesquisa básicos dos motores de pesquisa, de forma a obter a informação mais importante da mais recente. A actividade de monitorização torna-se mais pró activa se o indivíduo instalar canais ou *software* que encontrem e apresentem automaticamente informação baseada em palavras-chave ou assuntos.

Durante a pesquisa formal, é normal que inicialmente existam actividades de extracção, com alguma actividade de monitorização em complemento. A pesquisa formal faz uso de motores de pesquisa que possuam uma grande cobertura da *Web*, e que forneçam formas eficazes de procura e consequente extracção de informação. O indivíduo possui uma forte vontade de perder mais tempo na pesquisa, para aprender e usar os métodos complexos de pesquisa, e para avaliar as fontes encontradas ao nível da qualidade e da exactidão, porque não deseja perder informação importante. A pesquisa formal decorre em duas fases: pesquisa de multi-*sites* que identifiquem fontes relevantes seguida de uma pesquisa dentro do *site*. A extracção pode ser suportada pela actividade de monitorização, através de serviços como os de alerta de *sites* da *Web*, portais, e anúncios por e-mail para se poder manter actualizado com as últimas notícias.

III.5 Conclusão

A *World Wide Web* representa uma nova fonte de informação que é mais aberta, acessível e versátil que os seus predecessores. O alcance e riqueza da *Web* colocam desafios à gestão. A *Web* não é apenas um lugar onde uns procuram respostas a questões, mas também onde qualquer um pode pesquisar informação, fazer descobertas, e aprender mais sobre um determinado assunto. O estudo aqui apresentado centra-se

nas possibilidades das pessoas usarem a *Web* como apoio às suas actividades diárias e forneceu uma visão de como as pessoas usam actualmente a *Web*: indivíduos usam a *Web* como parte do seu trabalho em quatro modos de pesquisa de informação, variando do visionamento indirecto, visionamento condicional, à pesquisa informal e à formal. Em cada um destes modos de pesquisa de informação, os indivíduos detectam as necessidades de informação, adoptam estratégias de pesquisa, e usam a informação encontrada de formas distintas. Cada modo de pesquisa de informação na *Web* distingue-se pelas regulares sequências de acções do *software* de navegação da *Web* iniciadas pelo utilizador. Assim, o visionamento indirecto consiste em acções de iniciação e ligação; o visionamento condicionado caracteriza-se pela pesquisa, diferenciação e monitorização; a pesquisa informal é caracterizada pela extracção diferenciada e localizada; e a pesquisa formal consiste na extracção intensa e sistemática.

A um nível operacional pode-se assumir, com base nos modelos de pesquisa de informação na *Web* apresentados, que a pesquisa de informação dos utilizadores da *Web* aumenta à medida que evoluem nos quatro modos de pesquisa de informação. Na Tabela 5, apresentam-se algumas sugestões para a optimização da utilização da *Web*.

Modo do Uso da Web	Optimização da Utilização da Web
Visionamento Indirecto: iniciação e ligação	<ul style="list-style-type: none"> → Introduzir sistemas que possam pesquisar ou recomendar novos <i>sítes</i> ou <i>sítes</i> “parecidos com” → Motivar os indivíduos de um grupo a partilhar <i>bookmarks</i>, páginas Web, e URL’s → Criar portais cooperativos para suportar visionamento indirecto e accidental
Visionamento Condicionado: pesquisa, diferenciação e monitorização	<ul style="list-style-type: none"> → Treinar utilizadores a avaliar e a aumentar a prioridade ou a importância de informação → Tornar fácil aos utilizadores a partilha de informação baseada na Web via email ou através de fóruns <i>online</i> → Apresentar aos utilizadores serviços que lhes permitam subscrever e ser notificados sobre novos conteúdos de páginas Web
Pesquisa informal: diferenciação, monitorização e extracção	<ul style="list-style-type: none"> → Pré seleccionar fontes e motores de pesquisa de alta qualidade para tornar as pesquisas mais fáceis e informais → Pré-agrupar boas estratégias de pesquisa desenvolvidas por especialistas do assunto; permitir aos utilizadores visionar e aprender com essas estratégias → Educar os utilizadores em como avaliar a proveniência e a qualidade da informação
Pesquisa Formal: monitorização e extracção	<ul style="list-style-type: none"> → Educar os utilizadores sobre a abundância de fontes de informação que deveriam ser consideradas na pesquisa global: impressão, <i>online</i>, fontes humanas → Educar os utilizadores sobre quando usar: bases de dados <i>online</i> de serviços comerciais; biblioteca ou centro de recursos de informação; profissionais de informação → Treinar os utilizadores a utilizar técnicas avançadas de pesquisa: especificar ou generalizar a pesquisa; avaliar a exactidão e nova pesquisa; ligar “para a frente” e ligar “para trás”

Tabela 5 – Sugestões para otimizar a utilização da Web. Fonte: (Choo, 2000).

Capítulo IV – A Internet como Fonte de Informação Estatística

IV.1 Introdução

A informação estatística é cada vez mais considerada como um elemento indispensável à tomada de decisão a todos os níveis. Constitui uma das bases essenciais para a gestão moderna e assume, cada vez mais, uma importância estratégica no desenvolvimento económico e social das sociedades (Dias, 2002).

Informação estatística objectiva é vital para o desenvolvimento de uma sociedade livre e democrática e permite fundamentar solidamente decisões importantes, a todos os níveis. Segundo *Statistica Canada*¹⁶ a informação estatística pode ser usada para, por exemplo:

- Analisar o desempenho económico;
- Desenvolver políticas de desenvolvimento fiscal;
- Criar tarifas internacionais e negociações comerciais;
- Desenvolver políticas e programas que apoiem as pequenas empresas;
- Apoiar o desenvolvimento de políticas e programas de governo, no que se refere à economia e ao bem-estar da sociedade;
- Determinar os efeitos sociais e económicos da implementação de programas de governo;
- Apoiar as decisões legislativas do governo;
- Determinar a igualdade de pagamentos e transferências fiscais;
- Ajustar o índice de inflação, no que se refere a contratos e entidades;
- Desenvolver programas para promover competitividade nacional e internacional;
- Apoiar políticas e programas de imigração;
- Apoiar estratégias e programas no turismo;
- Avaliar o custo/eficiência de programas de saúde e de educação;
- Avaliar a eficiência do sistema da justiça;
- Seleccionar locais para construção de escolas.

¹⁶ Statistics Canada – <http://www.statcan.ca/english/about/used.htm> (consultado em 30/10/2004).

As estatísticas sempre acompanharam de perto a evolução das tecnologias da informação. Neste capítulo, pretende-se estudar a Internet como estratégia e política de difusão de informação estatística. Começa-se por ver como era divulgada a informação estatística antes da Internet e, em seguida, tentar perceber como a multiplicidade de meios de acesso à informação estatística, com recurso às novas tecnologias, estimulou a sua utilização.

IV.2 Informação Estatística Antes da Internet

Se a palavra “estatística” data do século XVII (Sousa, 1995), as estatísticas de massas que permitem conhecer os elementos da força do Estado, população, forças militares, riquezas, etc., vêm de há milhares de anos, tendo-se iniciado com os grandes impérios da antiguidade – para já não referir a simples contagem das coisas, dos fenómenos, dos casos, a qual, enquanto procedimento normal da vida quotidiana, remonta às mais antigas comunidades humanas, às origens da própria civilização.

Em Portugal, a introdução do vocabulário “estatística” revela-se bastante tardia, muito posterior às preocupações de natureza quantitativa desde cedo assumidas pelo Estado (Sousa, 1995). Com efeito, a produção de “estatísticas”, da informação baseada em números, surge na Baixa Idade Média e o termo estatística só aparece nos finais do século XVIII e apenas será divulgado nas memórias e discursos económico-políticos dos autores portugueses do primeiro quartel do século XIX.

Com a aproximação do Estado Liberal e a afirmação do conceito de Nação como base da administração, a cobertura estatística generalizada para o país começa a ser reclamada, pois o governo não se pode exercer eficazmente sobre o incerto, o desconhecido. Surgem planos para o cadastro do reino, levantam-se numeramentos de carácter sistemático, inicia-se a primeira grande série estatística sobre o comércio externo (Balança Geral do Comércio do Reino de Portugal, 1776-1831). Multiplicam-se os quadros estatísticos em diversas áreas da realidade social, apontam-se números globais, mas a informação é ainda, em grande parte, dispersa, recolhida em segunda mão, produzida por terceiros e nem sempre de acordo com os requisitos de qualidade e exigência que a estatística requer¹⁷.

A entrada na era estatística faz-se, portanto, gradualmente, ao longo do século XIX, com a criação de organismos que se fazem representar nos respectivos Congressos Internacionais. No século XX surge uma eficaz utilização dos dados recolhidos, com o desenvolvimento da Estatística como ramo aplicado da Matemática, ligando ao cálculo

¹⁷ Por exemplo, os dados da população eram solicitados aos párocos (Fernando Sousa, 1995).

das probabilidades, que vai permitir o fornecimento regular de indicadores de síntese, a perspectiva sequencial das tendências de desenvolvimento, a possibilidade de prospectiva; situação apenas possível com a criação do Instituto Nacional de Estatística (INE), em 1935, (Sousa, 1995).

Antes do surgimento da imprensa no século XV, a informação era transmitida oralmente ou através de manuscritos. Desde então, a informação tem sido criada em papel e gravada em vários formatos como livros, panfletos, relatórios, ficheiros, catálogos e jornais. A informação impressa pode ter várias vantagens para aqueles que procuram informação estatística: o formato é familiar, é transportável, e pode ser lida e pesquisada sem qualquer equipamento especializado (Large, Tedd & Hartley, 1999).

Durante o século XX a informação também era transmitida através de rádio, televisão, e, de uma forma crescente, através de meios electrónicos. Os desenvolvimentos de informação e das telecomunicações têm sido responsáveis pela explosão de informação: é reconhecido que a informação produzida entre 1970 e 2000 foi muito superior àquela que foi produzida nos 5000 anos anteriores (Large, Tedd & Hartley, 1999).

Na década de quarenta os computadores eram usados para processar informação numérica. A capacidade para armazenar informação textual desenvolveu-se na década de sessenta, com uma variedade de serviços. Estes serviços forneciam acesso a publicações que dariam aos cientistas informação sobre artigos publicados num jornal. Durante os anos oitenta, os serviços de pesquisa passaram a permitir o acesso a bases de dados de bibliotecas onde se poderia consultar artigos completos, jornais, enciclopédias e informação estatística. No final da década de oitenta, algumas bases de dados passaram a ser publicadas em Cd-Roms¹⁸. O Cd-Rom tornou-se, nos finais da década de noventa, um meio privilegiado para publicar informação cobrindo uma grande variedade de temas.

Contudo, durante os anos noventa, o surgimento de interface gráfico e o rápido desenvolvimento da Internet e da *World Wide Web*, transformou completamente a forma de armazenar e aceder a informação estatística. Hoje em dia, a possibilidade de encontrar informação estatística relevante na Internet de uma forma rápida e fácil – e muitas vezes, gratuita – é muito maior que encontrá-la num Cd-Rom. É um facto que torna a vida mais fácil a quem procura informação estatística, mas torna a vida um pouco mais complicada a quem pretende vender bases de dados com informação estatística em Cd-Rom (Jacsó, 1998).

¹⁸ Cd-Rom – CompactDisc-Read only memory

IV.3 Informação Estatística na Internet

Actualmente as informações estatísticas são obtidas, classificadas e armazenadas em meios magnéticos e disponibilizadas em diversos sistemas de informações abrangentes que fornecem aos pesquisadores/cidadãos e às organizações da sociedade, informações estatísticas “inteligentes” e necessárias ao desenvolvimento das suas actividades (Ferreira & Tavares, 2002). A expansão no processo de obtenção, armazenamento e disseminação de informações estatísticas, extensivamente facilitadas pelo uso dos recursos computacionais, tem sido acompanhada pelo rápido desenvolvimento de novas técnicas e metodologias de análise estatística de dados e de difusão.

No mundo, a informação estatística é produzida por uma grande variedade de institutos ou organizações: institutos de estatística, grupos de interesses específicos, empresas de seguros, empresas privadas, associações e organizações profissionais, investigadores independentes, instituições de ensino, organizações religiosas, grupos políticos, organizações sem fins lucrativos, e a imprensa. Alguma desta informação está disponível ao público em geral. A *World Wide Web* constitui o melhor meio de divulgação dessa informação. Contudo, no sentido de facilitar o acesso à informação estatística, é necessário criar um portal com uma interface fácil para o utilizador. Este portal tem de assegurar ao público em geral, assim como aos investigadores e estatísticos, notícia sobre a informação existente e acesso fácil e rápido à grande quantidade de informação presente nestes institutos (Ceaparu, 2003).

Segundo Hinchliffe & Duke (2001), a informação estatística:

- ➔ está disponível na Internet ou através de outras fontes; normalmente, as fontes governamentais de informação estatística são grátis para o público;
- ➔ pode estar disponível em bases de dados que exigem uma subscrição; em alguns casos, é necessário pagar uma taxa para que se possa ter acesso à informação desejada; as bibliotecas normalmente compram informação estatística para que os seus utilizadores tenham acesso de forma gratuita;
- ➔ está disponível apenas em formato impresso, no caso de dados especializados ou mais antigos; a informação estatística não está exclusivamente disponível em formato electrónico, pode estar disponível também em bibliotecas;
- ➔ pode aparecer em artigos de jornais ou revistas.

A informação estatística pode estar disponível na *World Wide Web* sob duas formas: conjunto de dados ou compilação de estatísticas (Hinchliffe & Duke, 2001). O conjunto de dados consiste em informação estatística na sua forma original, sem qualquer tipo de análise. A compilação de estatísticas consiste em informação analisada e transformada em informação utilizável. Este último tipo de informação estatística é apresentado, normalmente, sob a forma de tabelas, gráficos ou sumários. Em ambos os casos, é importante saber quem recolheu os dados, qual a metodologia utilizada e quando foram recolhidos.

Berinstein (1998) considera que a informação estatística é o tipo de informação mais difícil de encontrar. Os “pesquisadores de números” deparam-se com dois problemas: por um lado, para que a informação exista, alguém teve de a recolher, o que implica a existência de uma razão para o ter feito (sem razão ou fonte, não há informação estatística); e, por outro lado, a informação estatística está normalmente escondida em artigos, relatórios ou livros sem pistas anexadas que permitam que seja encontrada. Assim, Berinstein (1998) sugere cinco dicas de pesquisa para encontrar informação estatística na *World Wide Web*:

- ➔ **Dica de Pesquisa 1:** dar preferência às fontes constituídas, na sua maioria, por estatísticas como, por exemplo: almanaques, sumários estatísticos, bases de dados numéricas;
- ➔ **Dica de Pesquisa 2:** identificar ligações a áreas de estatísticas dentro de *sites* específicos; publicações, relatórios ou boletins; bibliotecas ou arquivos; bases de dados; e destaques da imprensa;
- ➔ **Dica de Pesquisa 3:** conhecer a estrutura da linguagem e a forma como a informação é exposta, pode trazer benefícios à pesquisa de informação estatística;
- ➔ **Dica de Pesquisa 4:** procurar informação estatística em forma de figuras, tabelas e gráficos.
- ➔ **Dica de Pesquisa 5:** usar termos específicos de pesquisa, tais como: estatísticas, estudos de mercado, números, previsão. Cada base de dados utiliza uma terminologia diferente, por isso é necessário perceber o que se procura.

Para além das dicas de pesquisa, Berinstein (1998) criou uma lista de fontes de informação estatística a que chamou “The Top Ten” dos Estados Unidos. As dez fontes de informação estatística na Internet são, na sua maioria, oficiais, uma vez que o governo

americano recolhe uma grande variedade de informação estatística para planear e apoiar as suas decisões políticas (Tabela 6).

Fonte	Descrição
1. Statistical Abstract of the U.S. http://www.census.gov/stat_abstract/	A informação disponibilizada consiste essencialmente em estatísticas (dica de pesquisa 1) e inclui dados de fontes privadas ou governamentais; abrange uma grande diversidade de temas, tais como: pessoas e sociedade, governo, saúde, indústria, etc.
2. U.S. Bureau of the Census http://www.census.gov	A informação estatística disponibilizada refere-se à população, habitação, comércio e vendas, salários, perfis locais, e indústria; este instituto não limita a informação aos dados dos censos, conduz regularmente inquéritos.
3. U.S. Bureau of Labor Statistics http://www.stats.bls.gov	A informação disponibilizada refere-se ao emprego, salários, preços, emprego no estrangeiro, e índices de preço e consumo (dica de pesquisa 2).
4. FEDSTATS http://www.fedstats.gov	Neste <i>site</i> , a informação estatística é pesquisada por mais de 100 institutos oficiais; aqui a pesquisa deve ser feita por assunto, ou através de um motor de pesquisa; no segundo caso, é aconselhável colocar palavras (por exemplo, tabela) no sentido de limitar a pesquisa (dica de pesquisa 4).
5. A Matter of Fact Disponível em bibliotecas e em CD-ROM	Consiste numa base de dados de factos estatísticos; as fontes incluem publicações alternativas e audiências do governo; informação estatística sobre economia, saúde, governo e ambiente.
6. World Almanac (HighBeam Library Research e Encarta Online) http://www.highbeam.com/library/index.asp http://encarta.msn.com/	A missão destas fontes é responder a questões colocadas pelo leitor comum.
7. American Demographics http://www.demographics.com	Esta revista disponibiliza informação estatística em texto completo; forte em informação relativa a demografia, a revista revela como as pessoas pensam, o que fazem, e o que procuram.
8. The CIA's World Factbook http://www.cia.gov/cia/publications/factbook/	É um compêndio de fácil utilização que traça o perfil de um país; informação sobre geografia, mapas e bandeiras, população, governo, economia, transportes, comunicação e defesa.
9. Major Business Journals (Forbes, Fortune, Business Week) http://www.forbes.com http://www.fortune.com http://www.businessweek.com/	Estas revistas de negócio precisam de abordar uma grande variedade de temas para chamar a atenção dos leitores; assim, possuem uma grande diversidade de informação (dica de pesquisa 3 e 4).
10. Newspapers o site http://www.esmonitor.com permite encontrar jornais de qualquer tema	Importantes notícias “numéricas” são dadas pelos jornais; pode-se encontrar, por exemplo, tendências tecnológicas e padrões populacionais, assim como resultados de inquéritos e estudos médicos (dica de pesquisa 3 e 5).

Tabela 6 – Top 10 de Fontes de Informação Estatística segundo Paula Berinstein. Fonte: (Berinstein, 1998)

Em Portugal, o Instituto Nacional de Estatística¹⁹, para além de desempenhar funções de serviço público estatístico, também assume um papel de grande relevo no mercado da informação, ultrapassando a condição de mero produtor de matéria-prima estatística, para oferecer, por sua iniciativa ou por encomenda, produtos e serviços de maior "valor acrescentado", nomeadamente estudos e análises nos domínios demográfico, social, económico e financeiro. A Biblioteca Digital de Estatísticas Oficiais disponibiliza no Arquivo Digital o acesso integral às publicações de informação estatística editadas pelo Instituto Nacional de Estatística até 2002 (1864-2002) e no Catálogo Bibliográfico a identificação de todas as publicações de âmbito estatístico de instituições nacionais, estrangeiras e internacionais disponíveis no Instituto Nacional de Estatística para consulta.

Para além do Instituto Nacional de Estatística, os vários Ministérios de Portugal produzem alguma informação estatística especializada na área de cada ministério. É o caso do Ministério das Finanças que possui a Direcção-Geral de Estudos e Previsão²⁰, que tem como tarefa principal realizar trabalhos de investigação científica no domínio da Economia, particularmente em matérias relevantes para o apoio da decisão e definição da política económica. Também é de referir o caso do Ministério da Ciência, Inovação e Ensino Superior que através do Observatório da Ciência e do Ensino Superior²¹ que tem a tarefa, de entre muitas outras, de assegurar o acesso, a recolha e o tratamento de informação estatística relativa aos sistemas científico e tecnológico e do ensino superior.

O Banco de Portugal²² também é uma fonte de informação estatística em Portugal. De acordo com a sua Lei Orgânica, o Banco de Portugal assegura a centralização e a elaboração das estatísticas monetárias e financeiras, da balança de pagamentos e cambiais.

Existem ainda as associações industriais ou profissionais, de carácter internacional como a *American Marketing Association* (AMA)²³ ou portuguesas, como a Associação Industrial Portuguesa (AIP)²⁴ que disponibilizam informações estatísticas sobre os sectores respectivos. Existe também informação com origem privada colocada em páginas *Web*. Os *sites* de grandes empresas de pesquisa de *marketing*, como a

¹⁹ <http://www.ine.pt/apresent/apresent.html> (consultado em 20/09/2004).

²⁰ Direcção-Geral de Estudos e Previsão do Ministério das Finanças: <http://www.dgep.pt/menuprinc.html> (consultado em 20/09/2004).

²¹ Ministério da Ciência, Inovação e Ensino Superior, Observatório da Ciência e do Ensino Superior: <http://www.oces.mces.pt/> (consultado em 20/09/2004).

²² Banco de Portugal: <http://www.bportugal.pt> (consultado em 20/09/2004).

²³ *American Marketing Association*: <http://www.marketingpower.com> (consultado em 7/02/2005).

²⁴ Associação Industrial Portuguesa: <http://www.aip.pt> (consultado em 7/02/2005).

*Fulcrum Analytics*²⁵, a *Forrester Research*²⁶ ou a *Marktest*²⁷ a nível nacional, que colocam *online* amostras de estatísticas e estudos para que os seus utilizadores se interessem pela aquisição dos relatórios completos.

IV.4 Confidencialidade/Segredo Estatístico

A sociedade dos nossos dias tende a depender cada vez mais da informação, o que implica o aparecimento de novos problemas sobre os direitos individuais e sobre a privacidade e confidencialidade de cada indivíduo. As pessoas, ao mesmo tempo que necessitam de informação sobre os mais variados temas, sentem-se preocupadas com o facto de essa informação estar a ser armazenada em base de dados e, a ser consultada por outras pessoas ou organizações desconhecidas para elas.

Os segredos, genericamente considerados e enquanto áreas de excepção ao princípio da liberdade de informação, existem nos mais diversos domínios. O elenco dos casos existentes vai desde a esfera pública à privada. Podem ainda distinguir-se entre segredos objectivos e subjectivos: os primeiros referem-se a qualquer pessoa em função de determinadas matérias; os segundos são derivados de determinada qualidade das pessoas que a ele estão sujeito. Na primeira categoria poderá incluir-se o segredo estatístico, enquanto que a segunda será o domínio privilegiado dos diversos segredos profissionais existentes (Conselho Superior de Estatística, 1999).

O segredo aplicado à estatística, pode dizer-se que assume a forma de ocultação ou reserva sobre as informações colhidas para esse efeito, seja por via directa, seja por via derivada com fundamento na tutela de determinados interesses. Entende-se, também, que o fluxo de informação estatística deverá ser, tanto quanto possível, um fluxo com apenas um sentido, sem a possibilidade de uma “viagem de regresso” ao seu ponto de origem²⁸.

Contudo, quando a informação estatística é desprovida de elementos singulares, ao ser conjugada com outra, ou da sua interacção com a realidade, pode, ainda assim, resultar a identificação singular da origem. Este processo varia em função da natureza da informação; por exemplo, não constituirá violação do segredo a divulgação de dados sobre indivíduos portadores de doenças contagiosas para o total do País. Não será assim em relação à mesma informação da freguesia X, onde existe, por hipótese, uma pessoa

²⁵ *Fulcrum Analytics*: <http://www.cyberdialque.com> (consultado em 7/02/2005).

²⁶ *Forrester Research*: <http://www.forrester.com> (consultado em 7/02/2005).

²⁷ *Marktest*: <http://www.marktest.pt> (consultado em 7/02/2005).

²⁸ Conselho Superior de Estatística (1999), *Análise e Reflexão sobre as Normas Actuais do Instituto do Segredo Estatístico*, Cap. I, N.º 2, Lisboa, INE.

nessa situação. Fácil será concluir que tal pessoa poderá ser, sem grandes custos, identificada, com todos os perigos daí emergentes, designadamente a discriminação.

Segundo o glossário de *Statistics Canada*²⁹, confidencialidade de informação estatística consiste na privacidade de informação fornecida pelos inquiridos, e na não divulgação de informações individuais dos mesmos.

Fienberg (1997) considera que a confidencialidade se refere, de um modo geral, à qualidade de informação estatística como uma obrigação de não transmitir essa informação a alguém não autorizado. A confidencialidade pode ser acordada entre indivíduos ou organizações, embora o direito à privacidade, no que se refere a informação estatística, normalmente tenha um significado para as organizações diferente daquele que tem para os indivíduos.

Nanopoulos & King (2002) referem que o princípio da confidencialidade estatística é efectivamente o contrato que liga o estatístico a todos aqueles que fornecem os seus dados individuais, quer seja voluntariamente (o que acontece frequentemente), quer por obrigação legal, com o objectivo de produzir informação estatística essencial à sociedade.

Quando a informação estatística é publicada, a sua confidencialidade tem de ser garantida. A informação estatística que é normalmente declarada confidencial encontra-se dividida em dois tipos: dados individuais e dados sob a forma de tabelas. No sentido de controlar a divulgação da informação, existem métodos que tentam tornar anónima a informação individual, quando esta é divulgada na forma de macrodata³⁰ (dados sob a forma de tabela) ou microdata³¹ (dados individuais).

Segundo Domingo-Ferrer e Mateo-Sanz (1999), estes métodos podem basear-se em duas áreas principais:

Restrição no Conjunto das Questões – método pelo qual as questões chegam à base de dados original de uma forma restrita, com respostas exactas ou negações. Esta abordagem possui cinco métodos: restringir o tamanho do conjunto das questões, controlar a sobreposição de questões, efectuar auditorias, restringir o acesso a células através da sua supressão ou partição (Fienberg,

²⁹ Glossário *Statistics Canada*: <http://www.statcan.ca/english/edu/power/glossary/gloss.htm> (consultado em 30-10-2004).

³⁰ Informação tabular é, normalmente, o produto do processo estatístico. As tabelas existem na forma de totais e a informação em cada célula é resultado de uma agregação de toda a informação individual relevante. (Eurostat, 1996).

³¹ Dados individuais com informação directa são aqueles cuja informação sobre a unidade estatística inclui o seu nome e/ou morada. Dados individuais com informação indirecta são aqueles cujo nome e/ou morada é desconhecida mas, informação adicional está disponível o que permite identificar o indivíduo num determinado tempo. (Eurostat, 1996).

1997). Os primeiros três métodos são direccionados a bases de dados disponíveis *online*, enquanto que os dois últimos são usados essencialmente para bases de dados *offline*;

Perturbação nos Dados ou nos Resultados – métodos em que as perturbações podem ser aplicadas directamente nos dados (*Data Perturbation*) ou apenas nos resultados (*Output Perturbation*). No primeiro caso, as bases de dados são alteradas, de uma forma sistemática ou aleatória, e as questões chegam apenas na base de dados alterada, e não na original. No segundo método, as questões chegam directamente à base de dados já restringida mas as respostas foram “perturbadas” de uma forma sistemática ou aleatória (Fienberg, 1997).

As tabelas constituem a forma mais comum e simples para representar informação estatística. Nanopoulos e King (2002) consideram que as tabelas podem ser reveladoras. Um importante aspecto das tabelas é que a divulgação apenas é possível se o assunto dos dados (de onde provêm) não for identificado directamente ou indirectamente. Como referido anteriormente, um dos métodos para proteger a informação consiste na supressão de células e formação de categorias com o objectivo de classificar variáveis. Por exemplo, uma tabela pode não ser reveladora para um leitor casual, mas poderá ser reveladora se o leitor tiver informação adicional relevante.

IV.4.1 Segredo Estatístico Português

A confidencialidade estatística é tratada na Lei da Estatística Portuguesa da seguinte forma:

- “1. O segredo estatístico visa salvaguardar a privacidade dos cidadãos, preservar a concorrência entre os agentes económicos e garantir a confiança dos informadores no sistema estatístico.*
- 2. Todas as informações estatísticas de carácter individual colhidas pelo Instituto Nacional de Estatística são de natureza confidencial, pelo que:*
 - a) Não podem ser discriminadamente insertas em quaisquer publicações ou fornecidas a quaisquer pessoas ou entidades, nem delas pode ser passada certidão;*
 - b) Constituem segredo profissional para todos os funcionários e agentes que delas tomem conhecimento;*
 - c) Nenhum serviço ou autoridade pode ordenar ou autorizar o seu exame.*
- 3. As informações individualizadas sobre pessoas singulares nunca podem ser divulgadas.*

4. Salvo disposição legal em contrário, as informações sobre a Administração Pública não estão abrangidas pelo segredo estatístico.

5. As informações sobre cooperativas, empresas públicas e privadas, instituições de crédito e outros agentes económicos não podem ser divulgadas, salvo autorização escrita dos respectivos representantes ou após autorização do Conselho Superior de Estatística, caso a caso, desde que estejam em causa as necessidades do planeamento e coordenação económica ou as relações económicas externas”³².

O princípio fundamental do Sistema Estatístico Nacional foi criado pela Base V da Lei n.º 1911, de 23 de Maio de 1935. Esta lei salvaguardava a privacidade individual, a defesa da concorrência entre os agentes económicos e a confiança dos cidadãos no sistema estatístico. No que se refere ao seu alcance, o princípio era tendencialmente absoluto, inclusive face ao poder judicial, admitindo-se excepções decorrentes do consentimento, da instrução e decisão de processos de contravenção estatística e de determinação expressa da lei³³.

Cinco décadas depois, a nova lei pretendeu instituir um regime mais liberal, objectivo, contudo ultrapassado pela transferência de competência para a sua libertação do Governo para o Conselho Superior de Estatística. A lei 6/89 de 15 de Abril mantém os vectores expressos em 1935: salvaguarda da privacidade, da concorrência entre agentes económicos e garantia de confiança no sistema. A análise e reflexão sobre as normas actuais do Instituto do Segredo Estatístico permitiu concluir que, ao nível do alcance, o leque de excepções foi ampliado e sujeitas à decisão casuística do Conselho Superior de Estatística, facto este que revela uma certa desgovernamentalização, mas que decorre de maior independência do órgão central do Sistema Estatístico Nacional. Vigora assim, como regra geral, a proibição de transmissão de informações de carácter individual referentes a entes privados. Para os entes públicos rege, pelo contrário, um princípio de transparência ou abertura, segundo o qual, salvo disposição em contrário, podem ser do conhecimento público todos os dados estatísticos, tendo em atenção o princípio e acesso à documentação administrativa que inspira o ordenamento do Conselho Superior de Estatística e o princípio da transparência que rege o sector público.

Segundo Assunção Cristas (2005), Presidente da Secção Permanente do Segredo Estatístico do Conselho Superior de Estatística, enquanto que o acesso a

³² Diário da República, I Série, N.º 88, Lei de Bases do Sistema Estatístico Nacional, Lei 6/89 de 15 de Abril, art. 5º.

³³ Conselho Superior de Estatística (1999), *Análise e Reflexão sobre as Normas Actuais do Instituto do Segredo Estatístico*, Cap. I, N.º 2, Lisboa, INE.

informações individualizadas sobre pessoas singulares é sempre proibido, os dados sobre cooperativas empresas públicas e privadas, instituições de crédito e outros agentes económicos, poderão excepcionalmente ser divulgados, assim o consinta o Conselho Superior de Estatística.

A referida presidente considera que, do panorama jurídico português, a privacidade das pessoas humanas é, assim, inviolável, que a Administração Pública deve ser transparente e que, em certos casos, a informação respeitante a determinadas pessoas colectivas é revelada. Se através de uma autorização pontual se consegue um fim contrário ao da lei – isto é, a divulgação generalizada da informação vedada – está-se a defraudar os intentos legais. Assim, para além da idade avançada da lei e de todo o rol de excepções, é chegado o momento de efectuar uma revisão não só do aspecto legal mas de toda a filosofia do Sistema estatístico Nacional (Cristas, 2005). É importante referir que a reformulação do segredo estatístico deve ser analisada nas três fases de acção: a da recolha de informação, a do tratamento de informação e da difusão/divulgação da informação. A última fase constitui o domínio de eleição do segredo estatístico. Nesta fase é necessário ter uma particular atenção, não só ao aspecto nominativo ou não nominativo da informação, mas também à sua interacção com a realidade. Informações só por si inocentes poderão produzir efeitos devastadores quando reassumem o contacto com a realidade.

IV.5 Levantamento de Sítios *Web*

A informação estatística encontra-se disponível através das mais variadas ferramentas, uma das quais se pretende analisar neste capítulo – a Internet. Conforme descrito no Capítulo II, informação deve constituir uma matéria-prima a partir da qual se constitua um produto útil à empresa (Carneiro, 1995). De modo a que se possa desenvolver o conhecimento do que se pretende significar com “matéria-prima”, alguns investigadores tentaram esboçar as diferenças entre dados, informação e conhecimento. Por exemplo, dados foram definidos como o resultado de observação directa de acontecimentos e informação como uma colecção estruturada de dados (Teskey, 1989). Os dados e a informação desempenham um papel importante no processo de decisão: informação consiste em dados de valor no processo de decisão que tem o potencial de transformar um estado de conhecimento (Belkin, 1975). Tal como é usada neste trabalho, informação refere-se a dados estruturados, que têm o potencial de transformar ou alterar um estado de conhecimento de indivíduo, assim como apoiar/fundamentar a tomada de decisões.

Assim, neste capítulo irá fazer-se uma descrição do Instituto Nacional da Holanda (*Statistics Netherlands*), do Canadá (*Statistics Canada*) e de Portugal (*Instituto Nacional de Estatística*). Em cada um dos institutos pretende-se saber qual a informação estatística que divulgam e, principalmente, como a fazem chegar, através dos seus sítios na *Web*, ao público em geral. Vai-se, também, apresentar uma instituição bancária, o Banco de Portugal, com especial incidência nos serviços de informação estatística.

IV.5.1 Caso *Statistics Netherlands* – StatLine

O Instituto de Estatística da Holanda (*Statistics Netherlands*) foi fundado há cerca de 100 anos com a missão de “compilar e publicar informação estatística incontestável, coerente, actualizada e que seja relevante para a justiça, política e investigação”³⁴. *Statistics Netherlands* fornece uma estrutura de informação como um serviço público – a informação produzida está disponível para todos ao mesmo tempo.

Statistics Netherlands compreende duas actividades primárias e duas secundárias. A nível primário, a produção de informação estatística e desenvolvimento de informação; a nível secundário, informação e comunicação tecnológica, incluindo desenvolvimento de metodologias e a gestão organizacional, com especial incidência na gestão de recursos humanos.

O processo de produção de informação estatística engloba a compilação, emparelhamento, processamento e análise de dados. A natureza destas actividades mudou ao longo dos anos de existência. O acesso a fontes administrativas, o conhecimento e capacidade para emparelhar e combinar os dados recolhidos e, posteriormente, transformá-los em conceitos estatísticos é um processo básico, enquanto que a confidencialidade dessa informação é uma condição essencial. *Statistics Netherlands* assegura, aquando da publicação de resultados de inquéritos, que nenhuma informação estatística referente às pessoas, empresas e instituições seja divulgada. Quando existe o cruzamento de base de dados, condições muito restritas são impostas para proteger a identidade individual.

No que diz respeito ao desenvolvimento de informação, esta actividade refere-se ao desenvolvimento de conceitos estatísticos, combinação e comparação de resultados através de sistemas consistentes de informação estatística e, ainda, a sua interpretação e respectiva publicação.

³⁴ *Statistics That Count*: <http://www.cbs.nl/en/organisation/corporate-information/index.htm> (consultado em 29/12/2004).

Embora a informação e comunicação tecnológica não constituam o objectivo principal de *Statistics Netherlands*, estas possuem uma importância vital para as suas actividades. A informação e comunicação tecnológica para além de constituírem uma infraestrutura física, também constituem uma rede computacional de ligação de dados com o mundo exterior.

Um outro aspecto com que o *Statistics Netherlands* se preocupa é a gestão organizacional, com especial interesse na gestão de recursos humanos. *Statistics Netherlands* pretende ser visto como uma instituição de alta qualidade e, para atingir esse objectivo, apostar nas habilitações dos recursos humanos é essencial.



Figura 10 – Homepage de *Statistics Netherlands*

IV.5.1.1 Serviços Disponíveis

A informação estatística publicada pelo *Statistics Netherlands* integra uma multitude de aspectos sociais, desde indicadores macro-económicos tais como o crescimento económico e preços de consumo, até rendimentos individuais ou de empresas. Os planos estatísticos deste Instituto são estabelecidos pela Comissão Central para a Estatística. Esta comissão fiscaliza a independência, imparcialidade, relevância, qualidade e continuidade do programa estatístico. O Director-Geral desta comissão tem a função de decidir autonomamente quais métodos a aplicar nos estudos, e se os resultados serão publicados ou não. *Statistics Netherlands* possui os seguintes serviços:

- *StatLine* – base de dados *online* e gratuita que permite pesquisar e seleccionar informação.
- *Key Figures* – principais indicadores relativos ao país: população, comércio e indústria, mão-de-obra, educação, rendimento, segurança nacional, macro-economia, mercado financeiro, governo e política, ambiente, desporto, justiça e saúde pública.
- *Population Counter* – Um contador de residentes na Holanda em constante actualização (Figura 11).
- *Just Published* – artigos e publicações recentes.
- *Press Releases* – comunicados à imprensa e que podem ser pesquisados por tema.
- *Library* – *Statistics Netherlands* publica os resultados dos inquéritos efectuados em livros, jornais, disquetes e Cd-Roms. A informação é pública, e nas duas bibliotecas existentes, qualquer pessoa pode consultar a informação estatística de que precisa. As bibliotecas possuem todas as publicações editadas pelo Instituto assim como por outros países e organizações internacionais tais como, a Organização para a Cooperação e o Desenvolvimento Económico (OCDE) e o Instituto Europeu de Estatística (Eurostat).
- *CEREM* – Centro para investigação de microdados económicos com o objectivo de permitir que investigadores externos executem pesquisas científicas sobre as empresas holandesas.

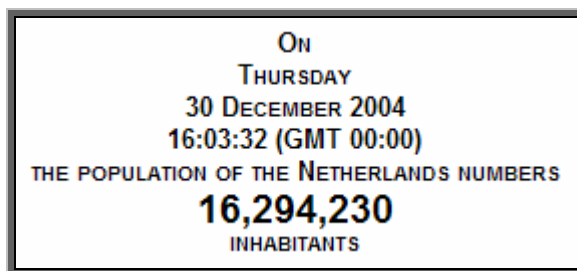




Figura 11 – Contador *Online* de População na Holanda

IV.5.1.2 Serviço *StatLine*

O serviço *StatLine* do *Statistics Netherlands* é um caso inovador no que se refere ao modelo de pesquisa de informação estatística pois baseia-se em quadros dinâmicos. A informação estatística é organizada em árvore, sendo o primeiro “nodo” constituído por um nível que é designado de *publicação*. O nível de *publicação* é por sua vez ramificado num conjunto hierárquico de dimensões e subdimensões, até chegar ao último “nodo”, no qual o utilizador selecciona então os itens que pretende visualizar e produzir o quadro estatístico pretendido.

StatLine consiste numa base de dados electrónica que contém informação estatística sobre assuntos sociais e económicos sob a forma de tabelas, gráficos e/ou mapas. Toda a informação contida no *StatLine* pode ser consultada, impressa ou importada sem qualquer custo associado.

A pesquisa de informação estatística pode ser feita através de dois métodos:

- Motor de pesquisa baseado em palavras-chave (Search ) – este método permite pesquisar através de palavras-chave e/ou de combinação de palavras-chave; permite ainda executar uma pesquisa avançada onde se poderá escolher o período temporal.
- Função “*Webselector*”  – *Webselector* é um programa que permite construir quadros simples ou com cruzamento de variáveis utilizando os dados do *Statistics Netherlands*. Este programa, fácil e inovador, funciona em 4 etapas:
 1. Seleccionar uma das *publicações* principais (Figura 12);
 2. Seleccionar um dos subtemas activados pela etapa anterior; nesta fase, também é activada a metainformação respectiva a cada subtema (Figura 13);

3. Seleccionar uma ou mais variáveis sobre a *publicação* escolhida (lado direito da janela) e respectivo espaço temporal, caso seja pretendido (Figura 14 e Figura 15).
4. Clicar em “*Show Data*” para gerar e visualizar o quadro pretendido (Figura 16).

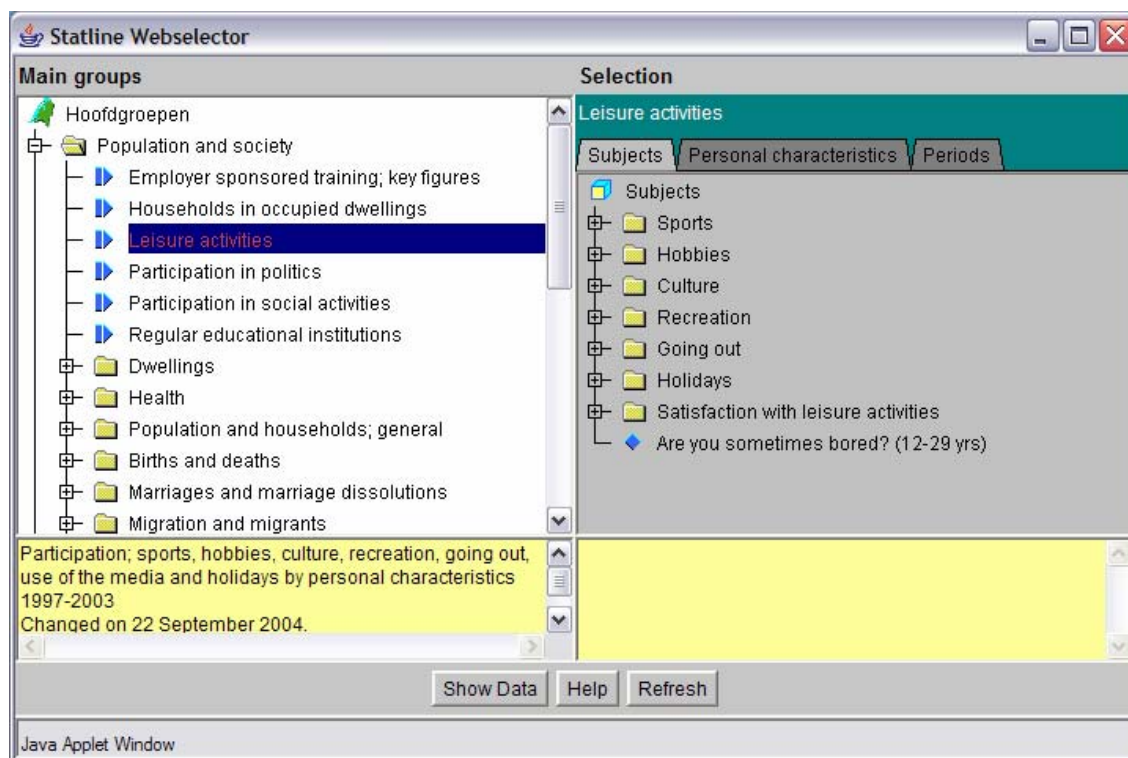


Figura 12 – Janela principal do programa *Webselector* – *StatLine*

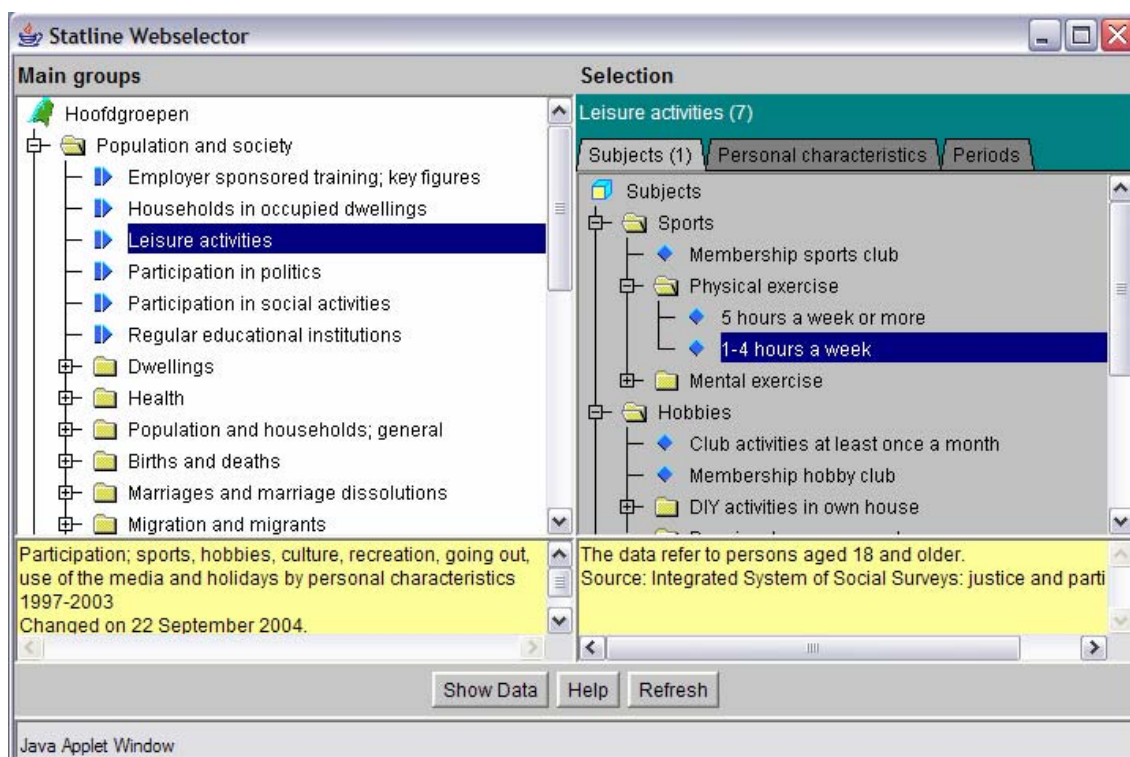


Figura 13 – Janela do programa *Webselector* com os sub-temas e metainformação activada

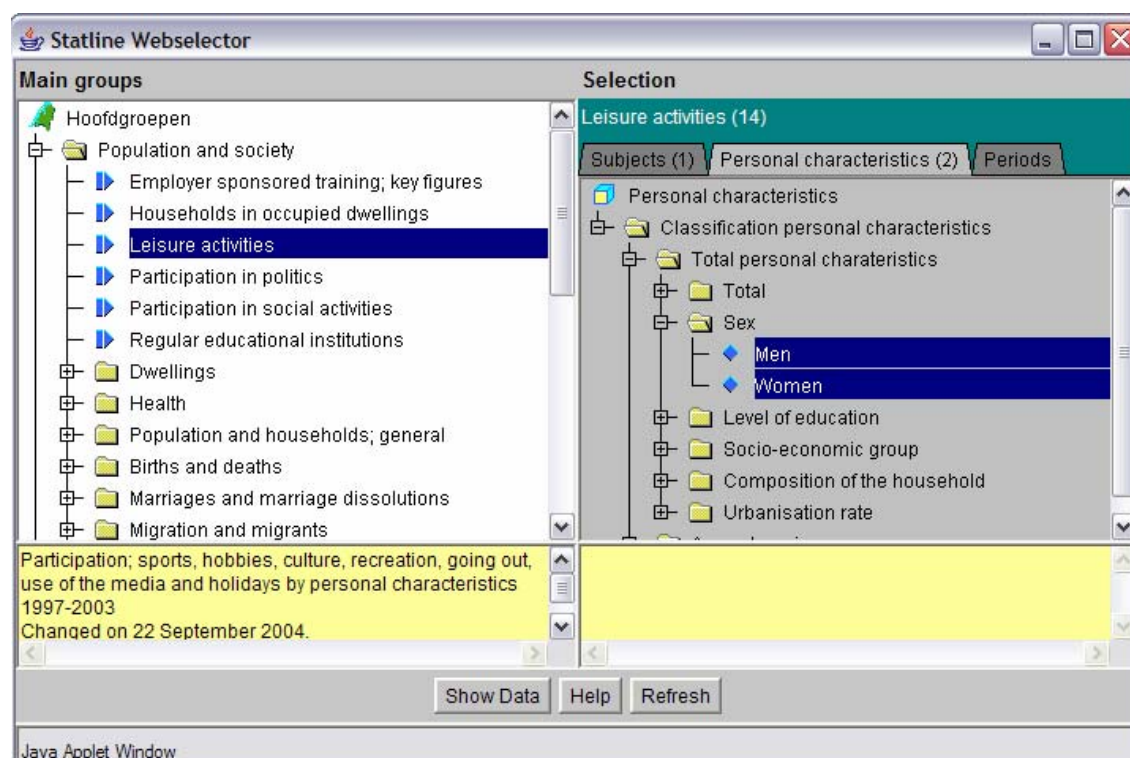


Figura 14 – Janela do programa *Webselector* com as variáveis específicas ao sub-tema escolhido

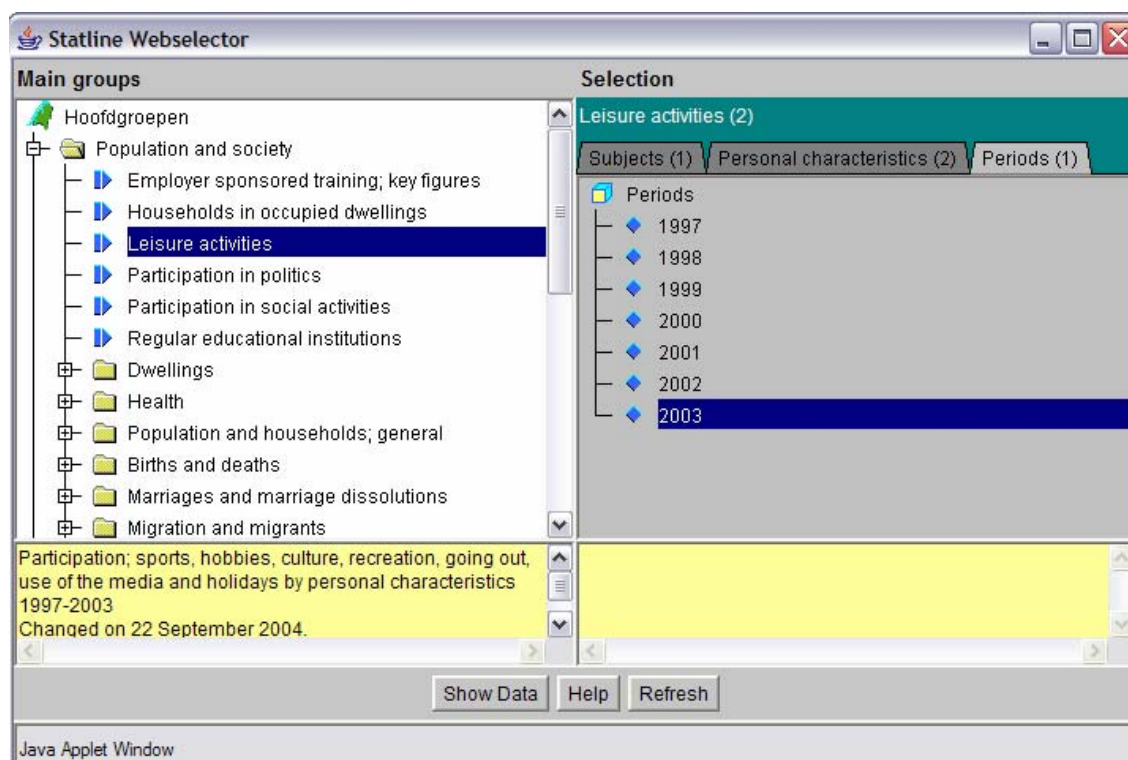


Figura 15 – Janela do Programa *Webselector* com a opção do período temporal



Figura 16 – Exemplo de quadro dinâmico gerado pelo programa *Webselector* – *StatLine*

A informação estatística produzida em formato de tabela pode ser apresentada de duas formas:


- ↪ *Matrizes* – Consistem em tabelas de duas ou mais dimensões. Por exemplo, uma matriz de informação estatística regional consiste numa tabela de três dimensões, em que a primeira contém tópicos relacionados com os municípios; a dimensão regional oferece a possibilidade de seleccionar regiões das quais se pretende informação; e, finalmente, a dimensão temporal permite escolher o ano ou o número de anos de que se pretende informação.
- ↪ *Tabelas* – Este tipo de *output* consiste em tabelas já previamente construídas tais como as que se podem encontrar em publicações estatísticas. É possível seleccionar um ou mais tabelas mas não é possível seleccionar a informação contida na tabela.

A informação estatística no *StatLine* é, como já foi referido, de acesso gratuito, mas o tamanho das tabelas que podem ser criadas têm um limite máximo: *online*, a tabela só poderá ter até 10 000 células, caso contrário terá de ser feito o *download* dos dados (até 100 000 células) com a possibilidade de o fazer em diferentes formatos: *Microsoft Excel*, *SPSS*, num ficheiro de texto, ou num ficheiro *html*. Nos casos em que os utilizadores precisem de grandes quantidades de dados, o Instituto disponibiliza a informação estatística em Cd-Rom.

Após ter sido formado o quadro dinâmico, o *StatLine* permite ainda manipular o seu aspecto, como por exemplo, recortar variáveis de forma a reduzir o número de células, ordenar e/ou transpor variáveis.

A informação estatística contida nas tabelas produzidas no *StatLine* pode também ser apresentada em formato gráfico, desde que sejam satisfeitas as seguintes condições:

- ↪ Tabelas com um máximo de 100 células;
- ↪ Tabelas cujas células possuam a mesma unidade;
- ↪ Tabelas que não tenham mais que duas dimensões com vários elementos.

Quando se visualiza o símbolo  associado à tabela construída, significa que esta satisfaz as três condições anteriores e é possível gerar um gráfico. Este ícone possui ainda uma ligação directa à página com o gráfico representativo da tabela. Caso as condições não sejam satisfeitas, o *StatLine* permite seleccionar uma parte da tabela e gerar, a partir dessa selecção, o respectivo gráfico. Qualquer gráfico gerado pode ser guardado e pode ser dos seguintes tipos: gráficos de barras, linhas e de áreas a duas dimensões (Figura 17).

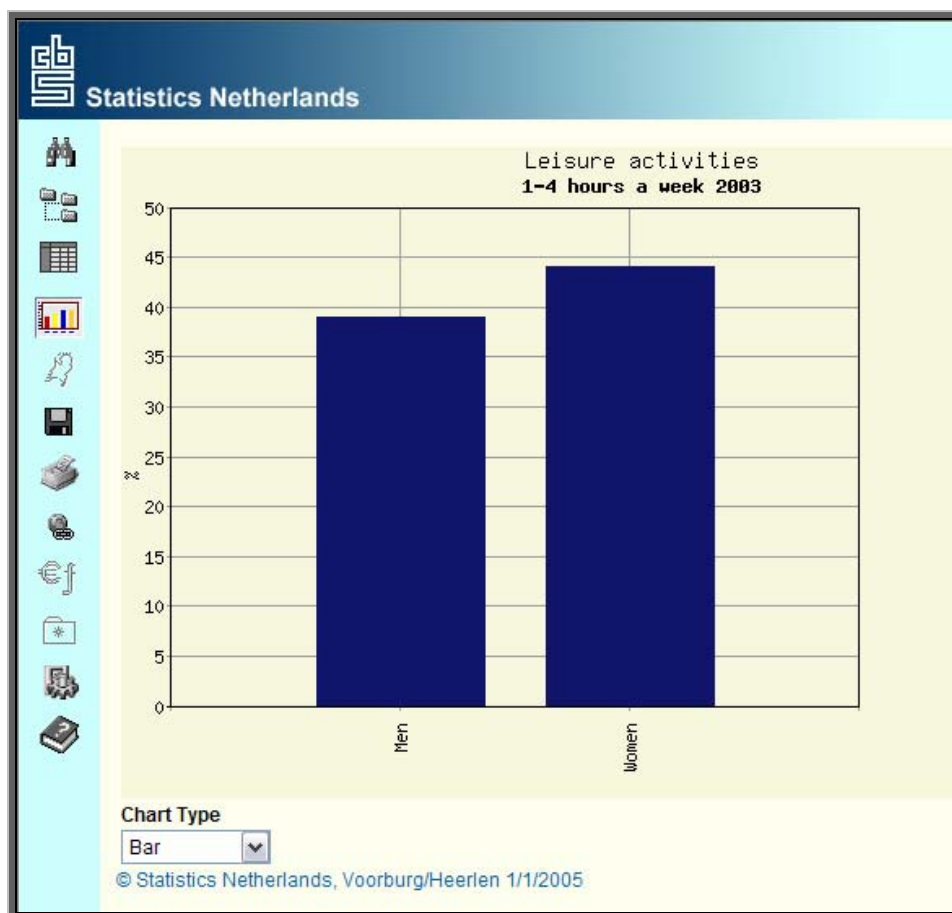


Figura 17 – Exemplo de um gráfico gerado no *StatLine*

Uma outra forma de apresentação de informação estatística possível no *StatLine* é o mapa estatístico. No caso das tabelas que contêm informação regional, é possível a construção de um mapa representativo dessa informação (Figura 18).

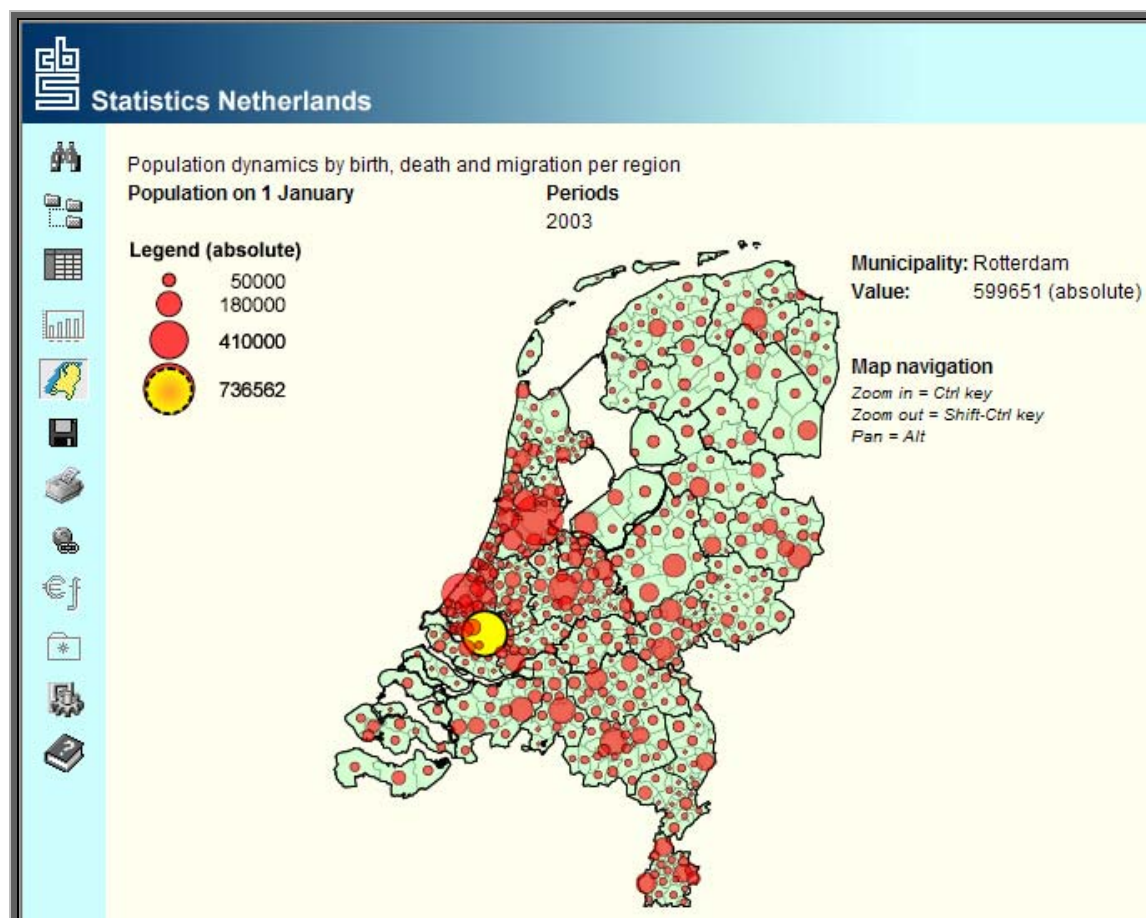


Figura 18 – Exemplo de um mapa estatístico gerado no StatLine

IV.5.2 Caso *Statistics Canada* – CANSIM³⁵

Statistics Canada tem a função de recolher, compilar, analisar e publicar informação estatística sobre a economia, instituições e/ou população canadiana. A sua missão – de fornecer informação objectiva e analisar a evolução da sociedade e economia do país – é fundamental para uma sociedade aberta e democrática.

A eficiência do *Statistics Canada* depende da sua credibilidade, da relevância da informação produzida e do seu acesso, em tempo útil, aos seus produtos e/ou serviços. Desde os anos noventa, a Internet tem exercido um papel estratégico e significativo na garantia da eficiência do Instituto. Neste aspecto, o objectivo do *Statistics Canada* é simples – fornecer informação estatística variada e detalhada sobre todas as características do Canadá através de um simples clique.

³⁵ CANSIM – Canadian Socio-Economic Information Management System, <http://cansim2.statcan.ca> (consultado em 3/01/2005).

No sentido de determinar as necessidades dos consumidores estatísticos, dos inquiridos e das empresas, *Statistics Canada* analisa o tráfego do seu *Website*, assim como conduz estudos de mercado e estudos de utilização da Internet. Todos os serviços e/ou produtos disponíveis quer *online*, quer *offline* são criados para complementar os canais de comunicação entre *Statistics Canada* e os seus clientes.



Figura 19 – Homepage de Statistics Canada

IV.5.2.1 Serviços Disponíveis

Statistics Canada produz informação estatística que permite aos canadianos conhecer e perceber melhor o país onde vivem – a sua população, recursos, economia, sociedade e cultura. No Canadá, a produção de informação estatística fornece uma fundamentação sólida para a tomada de decisões por parte de representantes políticos, de negócios, de coligações e de organizações sem fins lucrativos, assim como de indivíduos canadianos.

Segundo *Statistics Canada*, os canadianos avaliam a sua utilidade pelo conteúdo e relevância da informação estatística produzida. Por essa razão, a publicação da informação estatística produzida tem sido o serviço chave prestado pelo instituto. A Internet tornou-se o canal preferido dos consumidores para aceder às publicações estatísticas, e permitiu, ao Instituto, chegar a um número maior de canadianos que aquele atingido por meios de comunicação tradicionais.

Statistics Canada possui diversos serviços *online*, desde publicações a base de dados, censos e estudos. Desde 1995, o Instituto possui um *Website* que oferece um conjunto diversificado de informações estatísticas a um grande número de utilizadores:

- *Canadian Statistics* – consiste numa série de tabelas, actualizadas regularmente, relacionadas com a economia, com o país, e com a população.
- *Depository Services Program*³⁶ – programa criado em 1927 que consiste num depósito *online* de publicações estatísticas e que serve de ligação entre o governo federal canadiano e os seus clientes: o público canadiano, outros governos, universidades e empresas. De forma a atingir o objectivo principal – assegurar que os canadianos tenham igual acesso à informação estatística – este programa fornece publicações a uma rede constituída por mais de 790 bibliotecas no Canadá e a 146 instituições do mundo, a um custo reduzido.
- Um serviço de *e-commerce* que permite aos investigadores localizar e adquirir informação estatística através da base de dados *online* CANSIM.
- *Statistical Profile of Canadian Communities* – contém informação relativa a mais de 6 000 comunidades incluindo áreas metropolitanas, cidades, aldeias e colónias, desde os censos de 1996.
- *Data Liberation Initiative* – é um excelente exemplo de um método eficaz de melhorar os recursos de informações estatísticas a instituições de ensino. Para usufruir deste programa, as instituições interessadas pagam uma quota anual que permite o acesso ilimitado a vários ficheiros de microdados, bases de dados e publicações do *Statistica Canada*.
- *International Trade Database* – Base de dados de informação estatística relacionada com o comércio nacional e internacional que permite, por exemplo, encontrar mercados para uma possível exportação, determinar mercados competitivos e analisar o mercado canadiano ao nível das indústrias.

³⁶ Depository Services Program: <http://dsp-psd.pwgsc.gc.ca/index-e.html> (consultado em 3/01/2005).

- ↪ *Learning Resources* – apoia os programas educativos leccionados nas escolas básicas e secundárias. Ao nível dos estudantes são facultados dados e informação estatística de forma a permitir a realização das suas tarefas; ao nível dos professores são facultados dados, publicações e ferramentas de ensino criadas e seleccionadas por outros professores.
- ↪ *A Integrated Metadata Database* – fornece informação sobre todos os estudos e *outputs* divulgados pelo Instituto.
- ↪ *The Daily* – consiste no boletim oficial do Instituto editado desde 1932 e colocado *online* desde 1995. Neste boletim são publicadas as informações estatísticas recentes sobre a sociedade e/ou economia, assim como o anúncio de lançamento de novos produtos. *The Daily* é publicada *online* todos os dias úteis às oito e meia da manhã (hora canadiana oriental).
- ↪ *Infomat* – revista *online* que aborda as tendências sociais e económicas do país. *Infomat* está disponível, todas as semanas, em formato *html* e constitui uma fonte de um vasto conjunto de indicadores sócio-económicos.

Assim, o *website* é actualizado com informação estatística nova, todos os dias de manhã através do *The Daily*, do *Canadian Statistics*, e da base de dados *CANSIM*. Todos estes serviços combinados e, ainda, o módulo de *Community Profiles*, englobam 70% do tráfego do *website*³⁷.

Segundo um estudo do *Statistics Canada*, os seus clientes pertencem a uma grande diversidade de ocupações profissionais, das quais se salientam os governantes (cerca de 23%), estudantes (20%), empresas de prestação de serviços e contabilidade (15%), professores (13%), entre outros.

IV.5.2.2 Serviço CANSIM

CANSIM é a base de dados constituída por dados estatísticos em bruto fornecidos pelo *Statistics Canada* e que abrange um vasto leque de assuntos. Os dados são apresentados sob a forma de séries cronológicas: algumas séries são estáticas e apenas disponíveis para períodos históricos, mas muitas são extensivas a outros anos. As séries cronológicas podem ser apresentadas na forma de tabelas ou gráficos, que podem ser visualizados, impressos ou transferidos para o computador com ou sem um custo associado.

³⁷ <http://www.statcan.ca/english/about/online.htm> (consultado em 3/01/2005).

Em 2002, foi criada uma nova versão intitulada *CANSIM II* com a mesma estrutura e objectivos da anterior. Contudo, a ideia patente é de associar toda a informação estatística a tabelas de metainformação correctamente estruturadas e de manutenção automática. Esta base de dados possui milhares de matrizes sobre diversas áreas e centenas de milhares de séries que fornecem informação detalhada sobre os temas escolhidos. *CANSIM* possui informação estatística que cobre uma grande variedade de tópicos, tais como: agricultura, construção, ambiente, justiça, empresas, educação, governo, emprego, finanças, comunicação, energia, saúde, população, indústrias primárias e de serviços, comércio, demografia, transportes, preços, condições sociais, viagens e turismo.

A pesquisa de informação estatística pode ser realizada das seguintes formas:

- ➔ Por tema – todas as tabelas estão associadas a um ou mais temas a partir da lista geral do *Statistics Canada*. Quando se opta por este tipo de pesquisa, é-nos apresentada uma lista de dois níveis, em que o primeiro é constituído pelos 24 temas principais e o segundo, pelos subtemas respectivos (Figura 20).

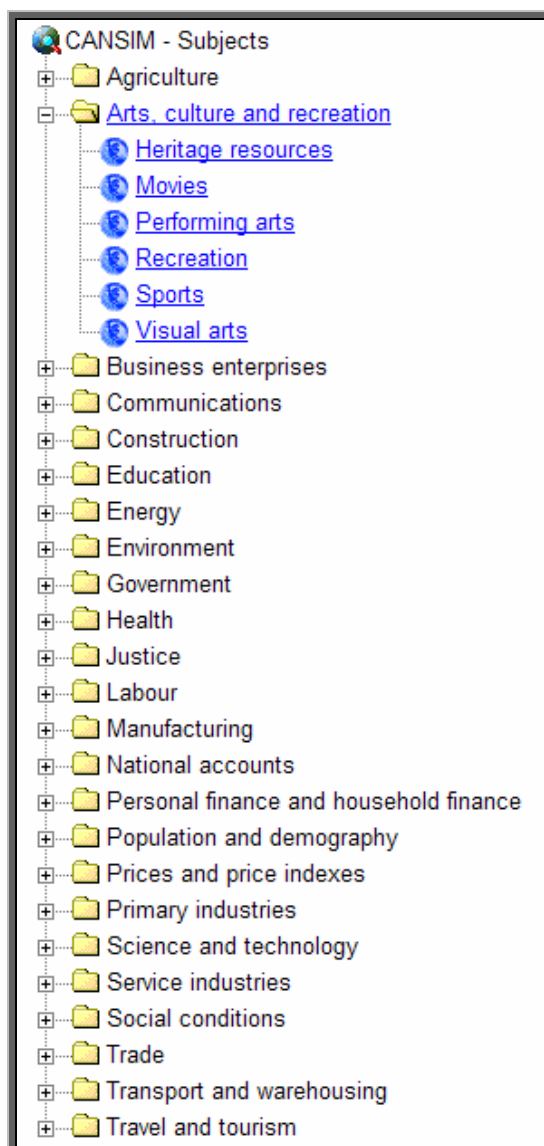


Figura 20 – Lista dos temas principais disponíveis na *CANSIM*, através da pesquisa por tema

➔ Por palavras-chave – a pesquisa pode ser feita através de palavras-chave ou frases que possam estar contidas nas tabelas. Esta pesquisa consiste em 4 etapas: primeiro, selecciona-se o tipo de procura (frase, todas as palavras, ou qualquer palavra); segundo, o tipo de tabela (todas as tabelas, tabelas activas ou não activas³⁸); terceiro, escolhe-se entre a procura por toda a tabela ou apenas ao seu título; e, por último, coloca-se a palavra-chave ou frase. É possível ainda optar por uma pesquisa avançada, quando se pretende utilizar comandos lógicos (e, ou, negação), (Figura 21).

³⁸ Uma tabela diz-se activa quando é constantemente actualizada; não activa, quando não é mais actualizada.

Figura 21 – Pesquisa por palavra-chave na CANSIM

- Por número da tabela – esta é uma das formas mais rápida de encontrar o pretendido, desde que já se conheça o número da tabela. CANSIM possui um directório com a lista das tabelas e o número respectivo (cada número de identificação possui a ligação para a tabela respectiva), (Figura 22).

Travel and tourism

Domestic travel

The Canadian Travel Survey (CTS) is a major source of data used to measure the size of Canada's tourism industry. Developed to measure the volume, characteristics and economic impact of domestic travel, it gathers data on more than 30 variables, including socio-economic characteristics, and expenditures.

- [426-0001](#) Canadian travel survey, domestic travel, by province and census metropolitan areas, annual
- [426-0002](#) Canadian travel survey, traveller characteristics, by province of destination, annual
- [426-0003](#) Canadian travel survey, travel characteristics, by province of destination, annual
- [426-0004](#) Canadian travel survey, travel characteristics, by province visited, annual
- [426-0005](#) Canadian travel survey, reallocated expenditures, by trip characteristics, annual
- [426-0006](#) Canadian travel survey, activity participation, by province of destination, annual
- [426-0007](#) Canadian travel survey, travel in Canada, by province of origin and destination and intraprovincial travel, bi-annual *Terminated*
- [426-0008](#) Canadian travel survey, travel in Canada, by age group, bi-annual *Terminated*

Figura 22 – Directório de tabelas da CANSIM

- Por número de série – consiste em fazer a pesquisa pelo número de série ou vector.
- Por número de estudo ou programa – A *Integrated Meta Data Base* (IMDB) contém informação sobre os estudos e programas estatísticos efectuados pelo *Statistics Canada*. A IMDB possui descrições de cerca de 450 estudos e

programas estatísticos identificados por um número. Após seleccionar o número, surge, para além da lista de tabelas criadas para esse estudo ou programa, as definições, fontes de dados e métodos utilizados para a construção das mesmas (Figura 23).



Figura 23 – Directório dos estudos e programas estatísticos realizados pelo *Statistics Canada*

Após a execução da pesquisa, por qualquer método, é apresentada a lista de tabelas correspondente. Esta lista é constituída, para além do tipo, número e título da tabela, por um sumário com informações sobre a tabela: frequência de actualização, unidade de medida utilizada, período temporal a que se refere e quais as variáveis que a constituem.

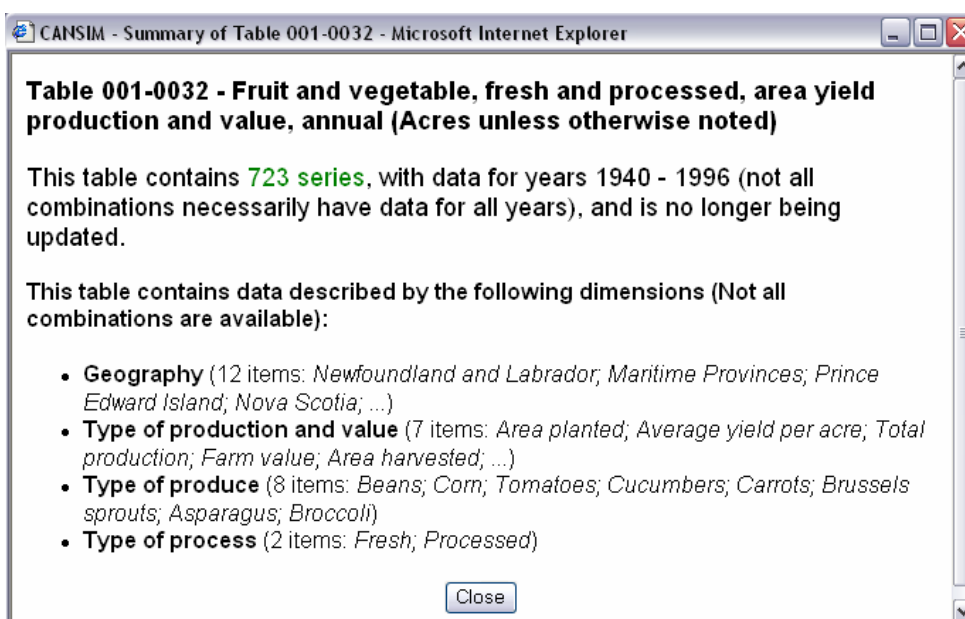


Figura 24 – Exemplo de um sumário de uma tabela na base de dados *CANSIM*

A informação fornecida pelo sumário permite escolher a tabela pretendida. Assim, surge uma lista de opções “*Table pick lists*” onde se escolhe a parte da tabela com a qual se pretende trabalhar (por exemplo, a localização e/ou período temporal) e as últimas novidades sobre o assunto.

Após a escolha da tabela, inicia-se o processo de compra, onde é facultado ao utilizador o custo total da(s) tabela(s) seleccionada(s). Nesta fase, o utilizador poderá optar por começar de novo a construção da tabela ou então avançar com a compra. Se o utilizador optar por comprar, terá de preencher um formulário com informações pessoais.

Table 127-0001 - Electric power statistics, monthly (Megawatt hour)							
Survey or program details: Monthly Electricity - 2151							
Unit of measure=Megawatt hour Geography=Canada							
Electric power, components	2001						
	January	February	March	April	May	June	July
Overall total generation¹	54,891,774	50,085,412	51,010,527	45,455,078	44,532,522	43,671,443	44,220,581
Total utility generation²	51,168,415	46,727,957	47,363,723	42,032,865	40,818,908	40,135,511	40,644,388
Total industrial generation³	3,723,359	3,357,455	3,646,804	3,422,213	3,713,614	3,535,932	3,576,193
Total hydro generation⁴	33,719,347	30,582,739	30,561,836	27,742,335	26,132,142	23,686,001	23,633,298

Source: Statistics Canada

Footnotes:

1. Total utility and industry generation made by hydro (including tidal and wind), steam (conventional and nuclear), internal combustion and combustion turbine.

2. A utility is an organization that, as its prime purpose, generates, transmits and / or distributes electric energy for resale. It can be private or public (provincial or municipal). Electricity can be made from hydro (including tidal and wind), steam (conventional and nuclear), internal combustion and combustion turbine sources.

3. Generation from establishments that are not part of the electric power generation, transmission and distribution industry (North American Industry Classification System (NAICS)): 22111).

4. Electric power generated from a plant in which the turbine generators are driven by falling water. This is the sum of utility hydro generation and industrial hydro generation.

Figura 25 – Exemplo de uma tabela obtida a partir da base de dados CANSIM

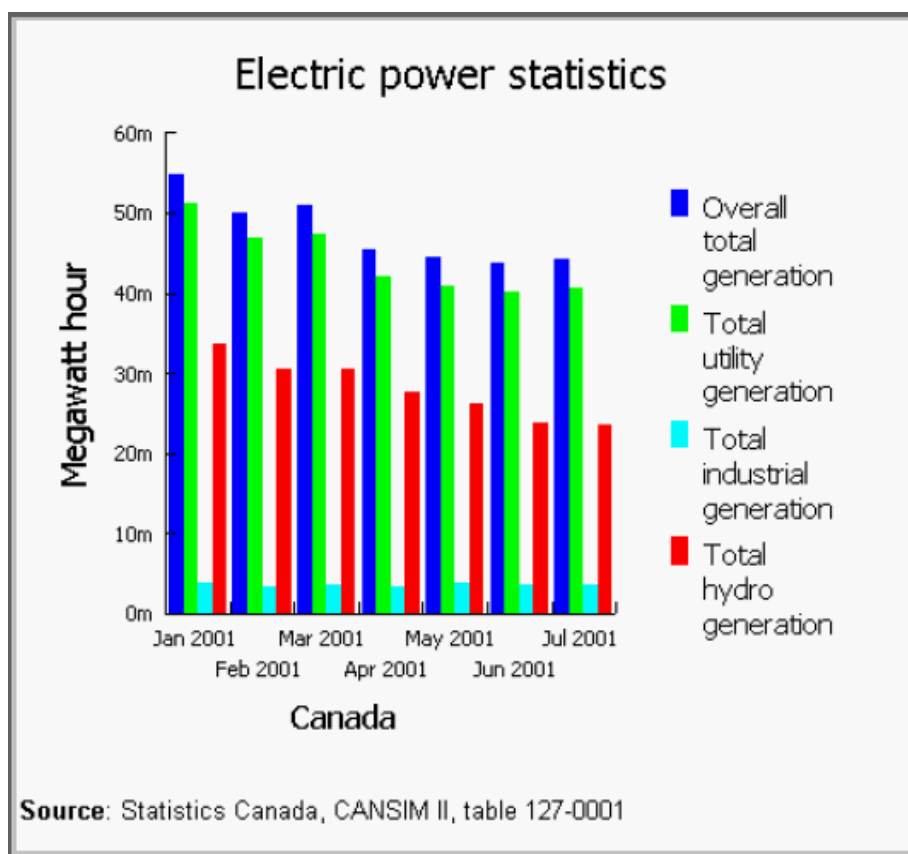


Figura 26 – Exemplo de um gráfico obtido a partir da base de dados CANSIM

IV.5.3 Caso Português

IV.5.3.1 Instituto Nacional de Estatística (INE)

A informação estatística permite reconhecer o ambiente económico e social em que se insere cada país, as necessidades e oportunidades que se revelam, e permite também explorar a tomada de decisão, quer a nível empresarial, quer a nível político, bem como ajudar a avaliar o grau de eficácia e o seu impacto. Como refere Paulo Gomes, em entrevista à “Vida Económica”, “o desenvolvimento da informação estatística gera novas oportunidades de gestão”³⁹. Torna-se, assim, fundamental que um país disponha de um Instituto de Estatística, que forneça informação não só para o público em geral, como também para empresas, estudantes, investigadores e órgãos de comunicação social. O INE foi criado em 1935 numa tentativa de dar resposta à procura cada vez maior desse tipo de informação.

³⁹ Sousa, João Luís, (2003), *Presidente do INE (Paulo Gomes) considera que as Empresas devem utilizar mais Informação Estatística*, *Vida Económica*, Ed. 18 Julho de 2003.

A criação do INE, para além de definir um marco histórico na Estatística Portuguesa, permitiu estabelecer pela primeira vez os princípios básicos da actividade estatística oficial, podendo agora falar-se de um verdadeiro Sistema Estatístico Nacional (SEN). Assim, foram estabelecidos os seguintes princípios:

- a centralização dos serviços;
- a autonomia técnica;
- a autoridade estatística;
- a responsabilidade dos informadores;
- o segredo profissional;
- a fiscalização das publicações estatísticas não directamente elaboradas pelo INE.

Em 1966 foi feita uma revisão que teve como finalidade actualizar e normalizar a legislação existente, bem como introduzir o princípio da coordenação estatística, sendo o novo Conselho Superior de Estatística a entidade responsável por velar pela sua observância e pela orientação superior do sistema. A criação de delegações regionais do Instituto foi prevista em 1973, aquando da reestruturação ao nível da organização dos seus serviços. Foi, contudo, em 1989 que se deu a mais profunda reestruturação do Sistema Estatístico Nacional e do INE, consubstanciada nos seguintes diplomas legais:

- Lei nº 6/89, de 15/4 – Bases Gerais do Sistema Estatístico Nacional;
- Decreto-Lei nº 280/89, de 23/8 – Estatutos do INE.

Em 1989 o INE passou a ser um instituto público, ao qual foi concedida personalidade jurídica, autonomia administrativa, financeira e património próprio.

O INE detém, em Portugal, a responsabilidade de assegurar a produção e difusão da informação estatística oficial. Continuando a desempenhar funções de serviço público estatístico, foi-lhe também aberta a possibilidade de assumir um papel de grande relevo no mercado da informação, ultrapassando a condição de mero produtor de informação estatística, para oferecer também, por sua iniciativa ou por encomenda, produtos e serviços, nomeadamente estudos e análises nos domínios demográfico, social, económico e financeiro. Tem como objectivo o exercício de funções tais como efectuar inquéritos, recenseamentos e outras operações estatísticas; criar, gerir e centralizar ficheiros de unidades estatísticas; aceder aos dados individuais (excepto dados de pessoas singulares) disponíveis nas entidades encarregadas da gestão de serviços públicos; realizar estudos de estatística pura e aplicada e proceder à análise económico-

social de dados estatísticos disponíveis; promover a formação de quadros do SEN e cooperar com organizações estatísticas estrangeiras.

Nos dias de hoje, o INE tem dezenas de publicações oficiais, não só em estudos demográficos, mas em diversos campos de aplicação, tal como indústria, comércio e educação.



Figura 27 – Homepage do Instituto Nacional de Estatística

IV.5.3.1.1 Serviços do INE

O INE possui uma orientação estratégica que assenta fundamentalmente nos princípios de satisfazer com produtos e serviços, adequados às necessidades dos utilizadores, segundo uma abordagem dinâmica e de favorecer a acessibilidade à informação. Assim, através do seu *website*⁴⁰ tem-se conhecimento dos seguintes produtos/serviços:

➔ Serviço INFOLINE – Serviço de Informação Online do INE;

⁴⁰ INE: <http://www.ine.pt> (consultado em 20/08/2004).

- ↳ Destaque – constitui o meio privilegiado de difusão de informação à Comunicação Social. Antecedendo a saída das publicações, são apresentadas no Destaque, com uma periodicidade diária, sínteses da informação produzida pelo INE. A disponibilização da informação a divulgar é registada num calendário de difusão semanal;
- ↳ Biblioteca Digital de Estatísticas Oficiais que disponibiliza no Arquivo Digital o acesso integral às publicações de informação estatística editadas pelo INE até 2000 (1864-2000) e no Catálogo Bibliográfico a identificação de todas as publicações de âmbito estatístico de instituições nacionais, estrangeiras e internacionais. A Biblioteca tem como objectivo promover a utilização e o conhecimento da informação estatística através das novas tecnologias, marcando uma nova etapa na difusão da informação estatística em Portugal;
- ↳ Catálogo *Online* de Produtos e Serviços que permite comprar publicações editadas pelo INE;
- ↳ Metainformação:
 - Conceitos Estatísticos⁴¹ – Versão *online* da base de conceitos, existente no INE, a qual permite assegurar a avaliação e a harmonização dos conceitos utilizados ao nível do Sistema Estatístico Nacional, condição necessária à coerência e compatibilização entre informação estatística de diferentes áreas temáticas. A versão disponível na Internet permite ao utilizador consultar e importar a totalidade dos conceitos. A pesquisa pode ser efectuada por palavra, área temática, fonte do conceito, código. Cada conceito estatístico dispõe de uma janela individual para acesso aos atributos que o caracterizam. Após a consulta é possível ao utilizador importar gratuitamente a informação obtida;
 - Classificações Estatísticas – constitui um dos mais importantes instrumentos para a harmonização e integração da informação estatística. As classificações apresentadas no INE estão organizadas por famílias de classificações afins, e por versões. Cada versão de uma classificação é um conjunto de categorias válidas num determinado intervalo de datas. Existem várias formas de pesquisa de classificações: por famílias, por ordem alfabética, por entidade responsável, entre outras;

⁴¹ Conceitos Estatísticos do INE: <http://conceitos.ine.pt/> (consultado em 2/02/2005).

- GEFSTAT, Gestão das Fontes Estatísticas⁴² – versão *online* da base existente no INE para a gestão dos inquéritos estatísticos, com vista a assegurar a coordenação técnica dos inquéritos realizados no âmbito do SEN. Para cada inquérito fica disponível a respectiva designação, a entidade responsável, as variáveis inquiridas e a metodologia adoptada para a recolha da informação. Para além disso, é ainda possível conhecer o historial dos inquéritos e a relação entre estes. A pesquisa a esta base de dados pode ser efectuada por âmbito de recolha, ano de validade, área estatística, designação, entidade responsável, método de recolha, número de registo, periodicidade, tipo de recolha, unidades inquiridas e variável inquirida. Para os inquéritos realizados a partir de 1997, é possível visualizar e imprimir os respectivos suportes de recolha de informação;
- ➔ Literacia Estatística – área destinada à publicação de textos estatísticos que permitam dar a conhecer as metodologias e metainformação utilizada na produção de informação estatística;
- ➔ *Revstat, Statistical Journal* – O INE edita, desde 2003, um revista semestral, publicada exclusivamente em inglês, que tem por objectivos a publicação de artigos de elevado conteúdo científico, o desenvolvimento de métodos estatísticos inovadores e a apresentação de investigação original, fundamentada em problemas reais. A REVSTAT cobre todos os ramos das Probabilidades e da Estatística, sendo também bem acolhidos estudos em importantes áreas de investigação nestes domínios;
- ➔ ALEA, Acção Local Estatística Aplicada⁴³ – este projecto constitui-se como um contributo para a elaboração de novos instrumentos de apoio ao ensino da Estatística para os alunos e professores do Ensino Básico e Secundário, com o objectivo de melhorar a literacia estatística;
- ➔ Webinq, Inquéritos do INE na *Web* – este serviço foi criado com o objectivo de melhorar o relacionamento dos informadores com o INE, permitindo a recolha de informação por via electrónica;
- ➔ Rede de Informação do INE em Bibliotecas do Ensino Superior – o público em geral e, em particular, estudantes, professores e investigadores dispõem de

⁴² Fontes Estatísticas do INE: <http://gefstat.ine.pt/> (consultado em 2/02/2005).

⁴³ ALEA: <http://alea-estp.ine.pt/> (consultado em 2/02/2005).

pontos de acesso em bibliotecas do ensino superior em todos os distritos do país.

O INE entende a informação estatística como um factor essencial para o desenvolvimento económico e social do país. Assim sendo, contribui activamente para a universalidade do acesso à informação, pretendendo responder da melhor forma possível às necessidades dos consumidores. Desta forma, é possível aceder a toda a informação estatística produzida pelo INE através de publicações, da Internet, de suportes informáticos ou consulta nas suas instalações.

O INE edita uma vasta gama de publicações estatísticas (cerca de 60 títulos), que podem ser adquiridos avulso ou por assinatura. Toda a informação sobre as publicações encontra-se referenciada no catálogo de publicações já referido. O INE proporciona, ainda, o acesso a importante parte da sua informação em suportes informáticos: bandas magnéticas, disquetes e Cd-Rom's.

A possibilidade de consulta (bem como importação) de toda a informação estatística publicada pelo INE é uma realidade através do seu *website*, onde se encontra o serviço INFOLINE. Igualmente elabora, sob encomenda, apuramentos estatísticos relativos a qualquer informação (publicada ou não).

Nas instalações do INE em Lisboa, Porto, Coimbra, Évora e Faro bem como no Serviço Regional de Estatística dos Açores e Direcção Regional de Estatística da Madeira, encontram-se disponíveis para consulta gratuita todas as obras editadas, parte da informação não publicada e ainda um grande número de outras publicações estatísticas nacionais e estrangeiras.

IV.5.3.1.2 Serviço INFOLINE do INE

Tendo acompanhado de perto a evolução das Tecnologias da Informação, não constitui surpresa que o INE iria utilizar a Internet para difusão de informação estatística. Em Maio de 1995, disponibiliza as suas primeiras páginas na Internet. Um ano depois, tem já o seu próprio domínio, com cerca de 50 páginas e uma média de 30 acessos diários (Dias, 1999). O *website* evolui e em 1997, a informação difundida é organizada em duas áreas: uma parte promocional e o serviço INFOLINE.

O INFOLINE é um serviço de acesso restrito, possibilitando a consulta e a importação de informação estatística do INE a utilizadores registados (Figura 28). Este serviço está direccionado para um público generalista, isto é, destina-se a todo o tipo de utilizadores de informação estatística. Numa primeira fase a informação a disponibilizar incluiu toda a informação estatística publicada (ou a publicar) durante o ano, isto é, uma

fotografia de todas as publicações editadas em papel pelo INE. A forma de apresentação representava uma imagem dos quadros estatísticos publicados (cerca de 5000) (Dias, 1999), somente diferindo destas pelas capacidades de organização e pesquisa (por temas e palavras-chave) que a Internet possibilitam.



Figura 28 – Formulário de entrada para utilizadores registados no INFOLINE

Conjuntamente com as melhorias dos meios tecnológicos são adicionados novos conteúdos bem como outras formas de aceder à informação estatística:

- ➔ Indicadores Estatísticos Gerais sobre os seguintes temas: Território e Ambiente; População e Condições Sociais; Economia e Finanças; Comércio Externo; Agricultura, Produção Animal, Silvicultura e Pescas; Indústria, Construção e Energia; Comércio Interno, Turismo e Outros Serviços. Estes indicadores são fornecidos em forma de tabelas estáticas, com possibilidade, gratuita, de fazer o *download* em formato Microsoft Excel (Figura 29).

INDICADORES ESTATÍSTICOS GERAIS

PRINCIPAIS : POR REGIÃO : POR TEMA

TEMA	Data de referência	Atualizado em
▲ Território e Ambiente		
Abastecimento e consumo de água	2002	10-02-2004
Drenagem e tratamento de águas residuais	2002	10-02-2004
Receitas e despesas dos municípios segundo os domínios de gestão e protecção do ambiente	2002	10-02-2004
Investimentos, custos e proveitos das empresas por sectores de actividade económica	2002	10-02-2004
Investimentos, custos e proveitos das empresas por domínios de ambiente	2002	10-02-2004
Despesas das administrações públicas em gestão e protecção do ambiente	2002	10-02-2004
Recolha e reciclagem de resíduos sólidos	2001	13-11-2003
Território e Ambiente - População, Área e Densidade Populacional	2001	14-01-2003
▼ População e Condições Sociais		
▼ Economia e Finanças		
▼ Comércio Externo		
▼ Agricultura, Produção Animal, Silvicultura e Pescas		
▼ Indústria, Construção e Energia		
▼ Comércio Interno, Turismo e Outros Serviços		

Figura 29 – Janela INFOLINE de acesso aos Indicadores Estatísticos Gerais

- Quadros Estatísticos para os temas referidos no serviço anterior e, ainda, para os temas Estatísticas Gerais e Sociedade da Informação e do Conhecimento. Estes quadros podem ser pesquisados por publicação ou por tema; é necessário um registo para efectuar o *download* respectivo ao quadro pretendido (Figura 30).



Figura 30 – Janela INFOLINE de acesso aos Quadros Estatísticos

- Séries Cronológicas para os temas já referidos nos Indicadores Estatísticos Gerais. As séries podem ser construídas com base na escolha do tema estatístico, do sub-tema estatístico e período temporal; é possível ainda visualizar o gráfico respectivo e inverter a disposição da tabela, disponível para utilizadores registados (Figura 31).

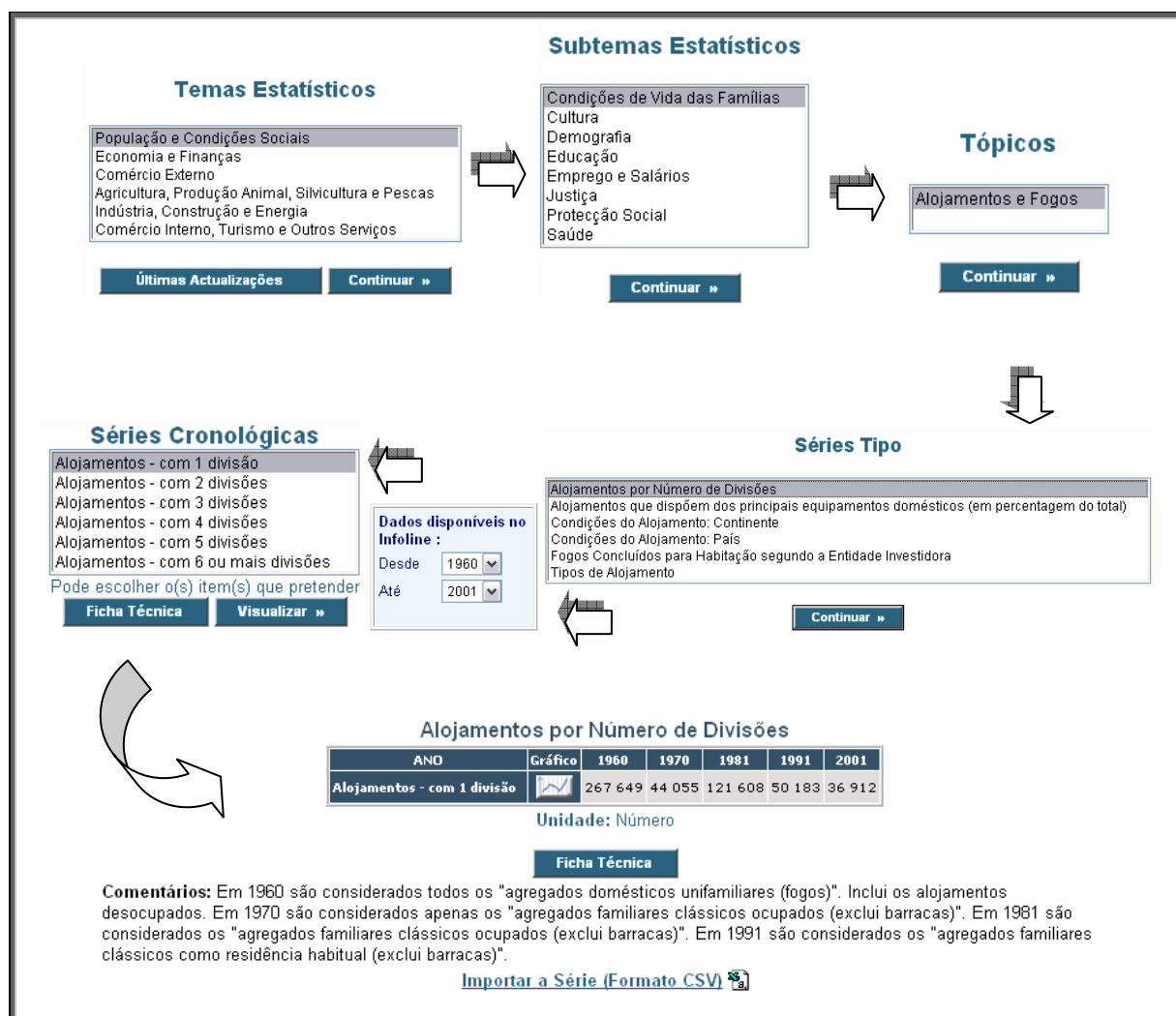


Figura 31 – Processo de obtenção de uma Série Cronológica no Serviço INFOLINE

- ➔ Pesquisa por unidade territorial que permite ao utilizador aceder a um vasto conjunto de indicadores que caracterizam uma determinada região, respeitantes aos diversos níveis geográficos (Figura 32).

Figura 32 – Janela de acesso à pesquisa por Unidade Territorial do serviço INFOLINE

- ➔ Folhas de Informação Rápida são constituídas pelos boletins mensais ou trimestrais, índices específicos, inquéritos mensais, entre outros; disponíveis em formato PDF⁴⁴ (Figura 33).

Figura 33 – Janela de acesso às Folhas de Informação Rápida do serviço INFOLINE

- ➔ Retratos Territoriais por NUTS e Concelhos/Distritos baseiam-se num sistema que permite a pesquisa de informação estatística relativa a um conjunto de indicadores fundamentais para os diversos níveis territoriais do país.
- ➔ Estudos – permite o acesso a estudos e análises estatísticas, editadas e divulgadas pelo INE em diversas publicações; estão organizados por área temática e por Revista, o que facilita a sua consulta (Figura 34).

⁴⁴ *Portable Document Format*, é um formato de ficheiro desenvolvido pela *Adobe System* para uso com o *Acrobat Reader*: <http://www.adobe.com> (consultado em 15/02/2005).



Figura 34 – Janela de acesso aos estudos no serviço INFOLINE

➔ Perfis de Utilizador – serviço realizado à medida dos interesses dos utilizadores. Para este serviço, o INE seleccionou e organizou a informação preexistente e apresenta informação nova, especialmente vocacionada para responder às necessidades de cada perfil de utilizador. Os perfis de utilizadores disponíveis são:

- Indicadores Estruturais – este serviço visa o acompanhamento dos progressos observados no conjunto dos Estados Membros da União Europeia nos domínios económico, social e ambiental relativamente à prossecução do objectivo estratégico definido no Conselho Europeu de Lisboa em Março de 2000 (Figura 35).

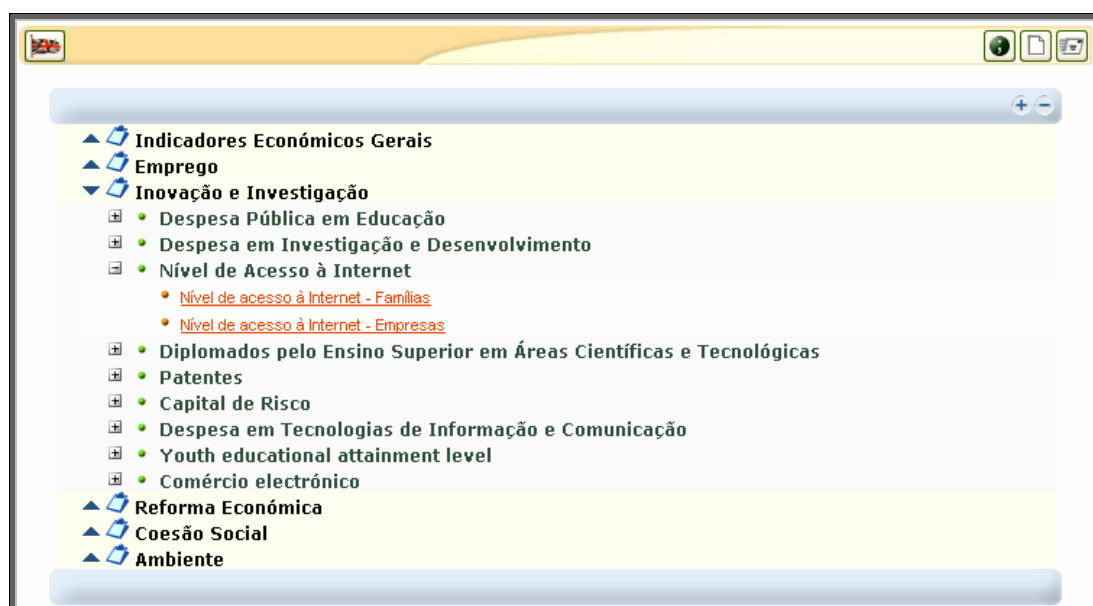


Figura 35 – Janela de acesso ao Perfil Indicadores Estruturais do Serviço INFOLINE

- **Perfil Género** – Este serviço visa proporcionar uma visão do peso relativo de cada um dos géneros (masculino e feminino) nas diversas áreas da sociedade portuguesa (Figura 36).

Áreas	Indicadores
População	Anual - Por Idade - Grupos quinquenais
Família	<input type="checkbox"/> Relação de masculinidade
Actividade, emprego e desemprego	<input type="checkbox"/> Relação de feminilidade
Educação e formação	Anual - Por Idade - Grupos quinquenais, Sexo
Conciliação trabalho/vida familiar	<input type="checkbox"/> População residente
Decisão	<input type="checkbox"/> População residente em percentagem
Saúde	Anual - Por Sexo
Crime e violência	<input type="checkbox"/> Índice de dependência de idosos
	<input type="checkbox"/> Índice de dependência de jovens
	<input type="checkbox"/> Índice de envelhecimento
	<input type="checkbox"/> Índice de longevidade
	<input type="checkbox"/> Índice de renovação da população em idade activa
	<input type="checkbox"/> Taxa de crescimento efectivo
	<input type="checkbox"/> População estrangeira com estatuto legal de residente
	<input type="checkbox"/> População estrangeira que solicitou estatuto de residente
	Anual - Por Sexo, Tipo de emigração
	<input type="checkbox"/> Emigrantes segundo o tipo de emigração
	Decenal - Por Idade - Grupos quinquenais, Sexo
	<input type="checkbox"/> Projeções da população residente
	Decenal - Por Sexo, Tipo de deficiência
	<input type="checkbox"/> População com deficiência

[Continuar](#)

Figura 36 – Janela de acesso ao Perfil Género do Serviço INFOLINE

- **Autarquias** – este serviço destina-se fundamentalmente aos agentes do poder local. Contém informação relativa aos diferentes projectos estatísticos que envolvem as Câmaras Municipais ou as Juntas de Freguesia, um quadro-resumo com os indicadores principais a nível concelhio e outra informação com especial relevo para as autarquias.
- **IntraStat, Estatísticas Correntes do Comércio Intracomunitário** – este serviço destina-se fundamentalmente aos responsáveis de informação estatística, contendo as informações necessárias à preparação e envio das declarações IntraStat, bem como a outros interessados nos procedimentos de recolha e nos resultados das Estatísticas do Comércio Intracomunitário.
- **Calçado** – Este perfil permite uma visibilidade mútua entre o organismo de produção de estatísticas oficiais portuguesas e as instituições representativas do sector do Calçado e destina-se a todos os empresários e responsáveis pelas áreas relacionadas com este sector.

Como resultado da nova estratégia de difusão de informação estatística do INE, desde Janeiro de 2002 que a visualização de toda a informação contida no INFOLINE, incluindo os documentos em formato PDF, é gratuita. Contudo, esta gratuitidade não dispensa a necessidade de registo no sistema. A partir desta data, o INFOLINE apresenta uma nova imagem e novas funcionalidades em resultado de melhoramento introduzidos, permitindo que a pesquisa de informação ocorra com maior rapidez e eficácia.

Com o intuito de atingir uma maior flexibilização do acesso à informação estatística, o serviço INFOLINE assenta toda a disponibilização da informação estatística em bases de dados de difusão e adequa esta difusão a diferentes perfis de utilizadores. Assim, a análise feita ao serviço permite salientar como pontos fortes a constante actualização de informação, o carácter único e oficial das estatísticas fornecidas, a visualização gratuita da informação, a facilidade de acesso e a flexibilidade de manuseamento da informação. Contudo, em comparação com serviços disponibilizados por outros institutos, verifica-se que o INFOLINE não tem ainda a possibilidade do utilizador definir os seus próprios cruzamentos, pois a informação a feio⁴⁵ não está disponível *online*, condicionando-o assim a escolhas pré-definidas. Mesmo assim, o INE consegue, de facto, estar mais perto do cidadão, ou seja, permite a que toda a população portuguesa tenha igual possibilidade de acesso à informação estatística que produz.

IV.5.3.2 Banco de Portugal

Neste ponto, pretende-se descrever o Banco de Portugal, nomeadamente a sua organização, as suas funções, responsabilidades e importância para o interesse nacional no que se refere à produção e divulgação de informação estatística. A escolha desta instituição, para ser analisada neste estudo, deve-se essencialmente ao facto de ter sido uma referência forte no questionário implementado e analisado na Parte II desta

⁴⁵ No INE existem dois tipos de informação: informação predefinida e informação a feio. A primeira consiste em informação publicada – informação disponível em publicações, independentemente do seu suporte, e em informação não publicada – informação pré-formatada e disponível em suporte papel e/ou electrónico. A segunda, informação a feio, e tendo em conta as limitações decorrentes do segredo estatístico e da protecção de dados pessoais, consiste em responder a pedidos de informação estatística específicos, os quais poderão ser feitos directamente aos centros de difusão de cada uma das Direcções Regionais do INE. A informação a feio é produzida com base nas seguintes bases de dados: Base Belém que contém dados físicos e económicos das empresas registadas em Portugal, de todos os sectores de actividade económica; Comércio Internacional que contém dados referentes às trocas comerciais com Estados Membros da União Europeia e com países terceiros; Bases Geográficas que contém informação geográfica da BGRE (Base Geográfica de Referenciação Espacial) construída a partir dos Censos de 1991 e da BGRI (Base Geográfica de Referenciação da Informação), construída a partir dos Censos 2001; Recenseamento Geral da População e Habitação e Recenseamento Geral Agrícola.

investigação. As empresas inquiridas referiram com frequência o Banco de Portugal como um outro *site* a que recorriam quando necessitavam de informação estatística.

O Banco de Portugal foi criado por decreto régio em 19 de Novembro de 1846, como consequência da fusão de duas instituições já existentes: o Banco de Lisboa, que funcionava como banco comercial e emissor desde a sua formação e a da Companhia Confiança Nacional, uma sociedade de investimento especializada no financiamento da dívida pública (Pereira, 1985).

A actividade do Banco de Portugal, hoje parte integrante do Sistema Europeu de Bancos Centrais, insere-se num contexto internacional e, em especial, europeu, caracterizado pela União Económica e Monetária⁴⁶. Além destes aspectos da inserção internacional da sua actividade, o Banco de Portugal participa nos trabalhos de um grande número de organismos internacionais, assegurando a representação internacional do Estado português junto de alguns deles.

De acordo com a respectiva Lei Orgânica⁴⁷, compete ao Banco "Gerir as disponibilidades externas do País ou outras que lhe estejam cometidas; agir como intermediário nas relações monetárias internacionais do Estado; velar pela estabilidade do sistema financeiro nacional, assegurando, com essa finalidade, designadamente a função de refinanciador de última instância; aconselhar o Governo nos domínios económico e financeiro, no âmbito das suas atribuições". Além das funções referidas e segundo a Lei Orgânica⁴⁸, compete ainda ao Banco de Portugal "a recolha e elaboração das estatísticas monetárias, financeiras, cambiais e da balança de pagamentos, designadamente no âmbito da sua colaboração com o Banco Central Europeu".

A gestão do Banco de Portugal efectua-se num contexto de grande autonomia do poder político, facto que se tem traduzido em maiores responsabilidades no domínio da prestação de contas e transparência. A sua organização é dirigida pelo Governador e pelo Conselho de Administração.

IV.5.3.2.1 Serviços ao nível de Informação Estatística

O Banco de Portugal necessita de informações estatísticas fiáveis para executar as funções que lhe foram atribuídas e para conduzir a política monetária e executar a política cambial. Por isso, compete-lhe assegurar a centralização e a elaboração das estatísticas monetárias, financeiras, cambiais e da balança de pagamentos. Nesse

⁴⁶ A União Económica e Monetária é uma zona com uma moeda única, que se traduz numa liberdade total de circulação de pessoas, bens, serviços e capitais.

⁴⁷ Lei Orgânica do Banco de Portugal, Capítulo IV, *Funções do Banco Central*, Secção I, Art. 12.º

⁴⁸ Lei Orgânica do Banco de Portugal, Capítulo IV, *Funções do Banco Central*, Secção I, Art. 13.º, 1.

sentido, o Banco de Portugal está autorizado por lei a exigir a qualquer entidade, pública ou privada, que lhe sejam fornecidas directamente as informações necessárias. Esta informação está sujeita ao regime legal sobre segredo estatístico e segredo bancário. O Banco de Portugal efectua as suas análises e destas resultam estudos, estimativas e previsões sobre a evolução da economia portuguesa, que o Banco de Portugal coloca à disposição dos decisores e agentes económicos, desempenhando uma importante missão de serviço público.

As estatísticas compiladas pelo Banco de Portugal, a sua análise, o seu estudo e as estimativas e previsões resultantes são agregadas em diversas publicações, objecto de uma significativa procura por parte dos decisores e agentes económicos, como se verifica no inquérito levado a cabo neste trabalho (Capítulo VI, Quadro 15). As publicações do Banco de Portugal são regulares, das quais se destacam:

- ➔ Relatório Anual – consiste num instrumento indispensável de análise sobre a economia portuguesa; é composto, normalmente por quatro capítulos: área do Euro onde é feita uma análise da actividade económica, monetária e financeira do Eurosistema, bem como do enquadramento externo; economia portuguesa onde é realizada a análise integrada da "Economia Portuguesa" ao longo do ano; actividade do Banco de Portugal e, por último, o balanço, contas, relatório e parecer do Conselho de Auditoria.
- ➔ Indicadores de Conjuntura – publicados mensalmente, na versão portuguesa e inglesa, integram um texto síntese, bem como gráficos e quadros relativos aos principais indicadores divulgados ao longo do mês; a versão electrónica permite visualizar e imprimir os ficheiros (em formato PDF).
- ➔ Boletim Estatístico – publicação mensal bilingue (português e inglês) disponibilizada em papel e em Cd-Rom, onde é divulgada a informação estatística compilada pelo Banco de Portugal, no âmbito das suas atribuições, nomeadamente as estatísticas monetárias e financeiras, as estatísticas da balança de pagamentos e da posição de investimento internacional e as estatísticas cambiais. Esta publicação inclui ainda capítulos específicos sobre principais indicadores, estatísticas das finanças públicas e estatísticas gerais da economia portuguesa. No *website* do Banco de Portugal é ainda disponibilizada uma versão electrónica, em formato PDF, para visualização no ecrã e/ou impressão

em papel, e uma versão, em formato csv⁴⁹, com as respectivas séries temporais.

- ➔ Boletim Económico – publicação trimestral, que integra artigos e textos sobre política e situação económica, bem como estudos aplicados sobre a Economia Portuguesa. As secções que o constituem são:
 - Discursos e intervenções públicas;
 - Textos de política e situação económica;
 - Artigos sobre Economia – a edição electrónica permite pesquisar, visualizar e imprimir os ficheiros correspondentes aos artigos publicados após 1996. A pesquisa pode ser feita por autor, por título e por ano;
 - Cronologia das principais medidas financeiras – actualizada mensalmente em simultâneo com os Indicadores de Conjuntura;
- ➔ Inquérito aos Bancos sobre o Mercado de Crédito – publicação trimestral que inclui o relatório síntese e os resultados detalhados para Portugal do inquérito desenvolvido pelo Eurosistema para a área do Euro; disponível apenas *online* e formato PDF.
- ➔ Cadernos do Banco de Portugal – pequenas publicações cujo objectivo principal consiste em informar e esclarecer os consumidores de produtos financeiros, das características das operações à sua disposição, do seu modo de funcionamento e dos direitos e deveres que lhes assistem. Os cadernos são de pequeno formato, estruturados sob a forma de perguntas e respostas e agrupados por temas bem identificados.
- ➔ Séries Longas para a Economia Portuguesa – nesta publicação apresentam-se estimativas de séries estatísticas e respectivas notas metodológicas, para um conjunto de variáveis económicas e para um período relativamente longo, desde finais da década de 40 ou início da década de 50, consoante as variáveis. Foram cinco as vertentes temáticas consideradas: Estatísticas Monetárias e Financeiras 1947-1995; Estatísticas de Balança de Pagamentos 1948-1995; Contas do Sector Público 1947-1995; Estatísticas da População, Emprego e Desemprego 1953-1995; Estatísticas da Produção, Despesa e Rendimento 1953-1995.

⁴⁹ Csv (*Comma Separated Values*) – ficheiro de texto cujos campos são divididos por ponto e vírgula normalmente utilizados para importar dados em bases de dados e folhas de cálculo.

- Quadros de Situação Sectoriais – consiste num programa para ambiente *Windows* que permite a consulta, impressão e exportação dos dados que figuram nos Quadros de Situação Sectoriais, editados pela Central de Balanços⁵⁰ do Banco de Portugal (Figura 37). Os dados são recolhidos através do preenchimento de um inquérito anual pelas empresas, em papel ou em disquete, e, depois de convenientemente analisados e validados, dão origem a "Quadros-Síntese" individuais e de centralizações globais. A partir destas últimas são editados "Quadros de Situação Sectoriais" por sectores de actividade económica, com o máximo de desagregação possível (Figura 38). Os Quadros de Situação Sectoriais, sem valores individuais, têm sido editados em suporte papel ou em Cd-Rom. A partir do exercício de 1999, os Quadros de Situação Sectoriais passaram a ser distribuídos gratuitamente através da página do banco na Internet. Neste momento estão disponíveis para os anos de 1991 a 2002.



Figura 37 – Programa QSS que permite a consulta, impressão e exportação dos dados que figuram nos Quadros de Situação Sectoriais, Banco de Portugal

⁵⁰ A Central de Balanços reúne um importante conjunto de informações económicas e financeiras das empresas que voluntariamente com ela aceitam colaborar. Além de dotar o Banco de Portugal de instrumentos válidos para a análise da situação e das perspectivas económico-financeiras dos sectores de actividade, a Central de Balanços proporciona aos empresários portugueses informações úteis para a gestão e enquadramento das suas empresas no respectivo sector de actividade e faculta às Associações Empresariais, Universidades, Consultores, Analistas e outros utilizadores dados agregados para apoio na realização de estudos.

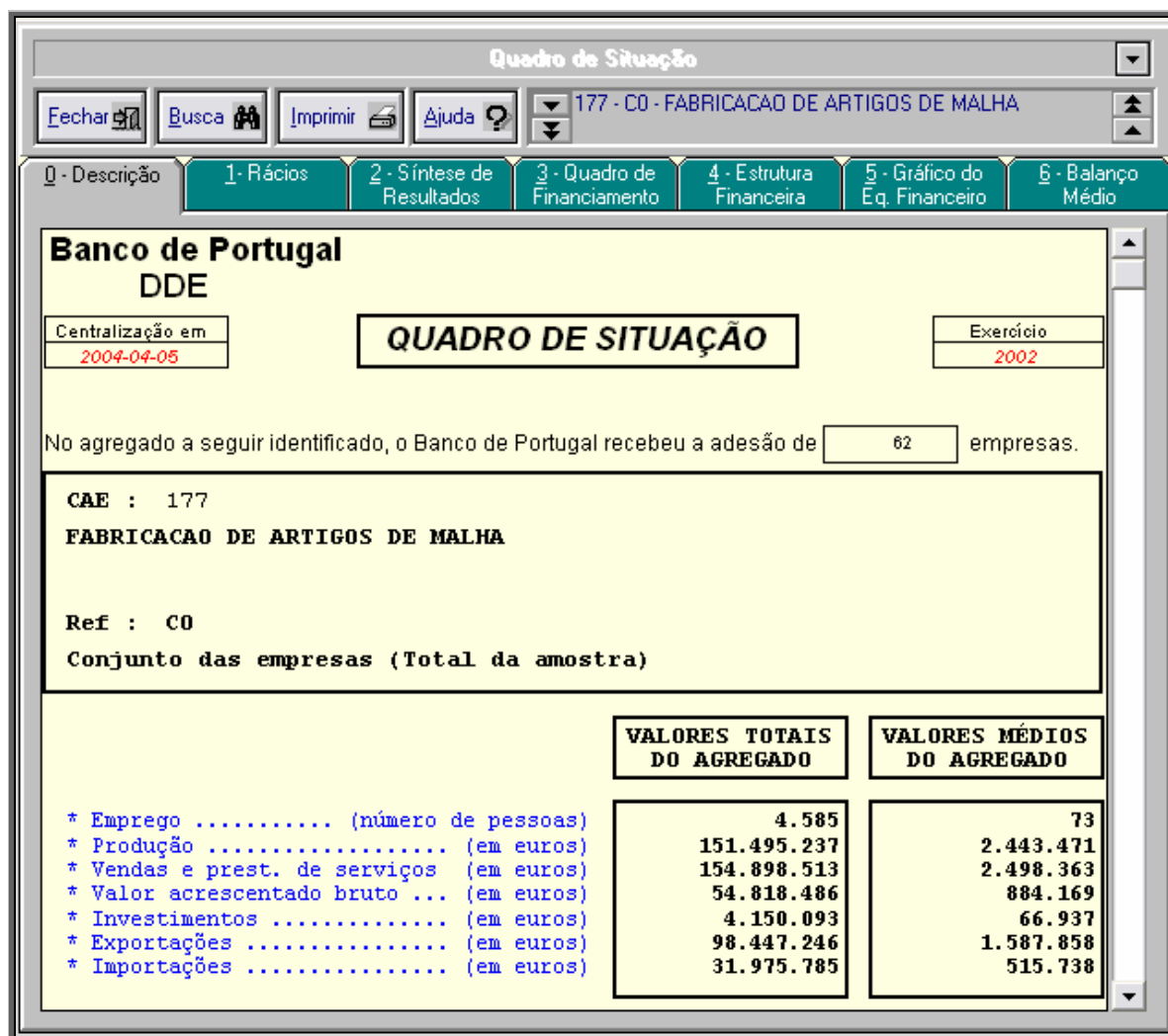


Figura 38 – Exemplo de um Quadro de Situação Sectorial obtido através do programa QSS

Como foi referido atrás, alguma informação estatística é obtida através da colaboração das empresas. Para isso, o Banco de Portugal possui as seguintes aplicações de recolha de informação estatística:

- QIPE2003 – Questionário ao investimento directo de Portugal no exterior;
- ARINTO – Aplicação para recolha de informação de títulos;
- IMCD – Programa para recolha/validação dos dados do inquérito à actividade nos mercados de câmbios e de produtos derivados;
- ICB – Aplicação de recolha do inquérito anual da central de balanços.

Como refere Vítor Constâncio, Governador do Banco de Portugal⁵¹, “os cidadãos esclarecidos e actantes fazem escolhas adequadas, utilizam instrumentos racionais e

⁵¹ <http://www.bportugal.pt> (consultado em 20/09/2004).

melhoram as relações estabelecidas entre os diversos agentes económicos intervenientes e entre estes e as autoridades de supervisão”. Acrescenta ainda que, “é fundamental que à informação disponibilizada se aceda facilmente e que a linguagem utilizada seja clara e compreensível pelo grande público”. De facto o Banco de Portugal consegue atingir os seus objectivos, pela publicação e divulgação de todo o tipo de informação, estatística ou não, através dos seus boletins e/ou cadernos, em formato electrónico, Cd-Rom e/ou em papel. Desta forma, o Banco de Portugal consegue promover um maior conhecimento e sofisticação no uso dos produtos financeiros, essenciais para a tomada eficiente de decisões ao nível de poupança e/ou investimento.

IV.6 Conclusão

Os serviços de divulgação de informação estatística apresentados baseiam-se em bases de dados: StatLine da Holanda, CANSIM do Canadá e INFOLINE de Portugal. Cooke (1999) define base de dados como um conjunto de registos, que por sua vez contém detalhes de diferentes dados, sejam numéricos, textos ou imagens, e que normalmente estão disponíveis através de pesquisa. Como se viu, existem algumas fontes de informação estatística através da Internet que podem ser descritas como bases de dados e que foram desenvolvidas especificamente para uso via WWW.

Cooke (1999) apresenta uma lista de critérios (Anexo 1) para comparar e avaliar bases de dados disponíveis na *World Wide Web*. Com base em alguns destes critérios e na descrição feita aos serviços referidos será estabelecida uma comparação entre estes.

Os serviços ao nível da divulgação de informação estatística do Banco de Portugal não são comparáveis com os dos três institutos analisados, pelo facto de este não ter como actividade principal a produção e difusão de informação estatística.

☒ Qual o objectivo pretendido, cobertura e audiência da base de dados? Esta informação está disponível em qualquer outra informação introdutória ou ficheiros de ajuda?

Os futuros utilizadores de qualquer base de dados devem, em primeiro lugar, ter acesso a um guião de apoio à sua utilização. É importante saber quais as potencialidades de pesquisa e, principalmente qual a sua cobertura ao nível de áreas temáticas. Este é, de facto, um ponto forte da base de dados CANSIM. No *website* do instituto canadiano é possível fazer *download* de um guião de utilizador (em formato pdf), que fornece todas as indicações para consultar a base de dados e, ainda, aceder a um tutorial animado que fornece exemplos das várias formas de pesquisa possíveis. No que se refere às áreas

temáticas, quando se acede à pesquisa por tema tem-se acesso à listagem de temas de uma forma clara e objectiva (lista apresentada em forma de árvore).

A base de dados StatLine possui igualmente um guião em formato *html* onde o Instituto dá a conhecer aos utilizadores as diferentes formas de pesquisa, assim como os formatos de *download* de ficheiros. Relativamente às áreas temáticas, estas são mostradas quando se opta pela função de pesquisa “*Webselector*”, que pela facilidade de utilização, permite o acesso a todos os temas e subtemas de informação estatística disponíveis.

No caso do Instituto português, este critério não é satisfeito uma vez que não existe nenhum guião que explique o funcionamento do serviço INFOLINE. Contudo o *website* possui um “mapa do site” onde os utilizadores poderão ter acesso à apresentação do Instituto e aos seus produtos e serviços. No caso do INE, o *website* possui, na página de entrada, um *pick and go*, com todos os temas de informação estatística, constituindo assim uma vantagem em relação aos outros institutos, em que apenas após a escolha de uma pesquisa se tem acesso à listagem de temas.

Na avaliação deste critério, as bases de dados CANSIM e StatLine apresentam vantagens em relação ao INFOLINE.

☒ Qual é a cobertura da base de dados? Que áreas temáticas e materiais são abordados? A base de dados está completa numa determinada área? Qual é a cobertura retrospectiva da base de dados? A informação sobre o material que está incluído na base de dados está disponível? Existem ligações a outras possíveis fontes de informação?

O nível de detalhe de cada registo constitui um factor importante a ser considerado numa base de dados. É importante ainda, saber o que é fornecido em cada registo, isto é, por exemplo, o que significa um determinado valor, que unidade é representada, a origem dos dados, entre outros.

A base de dados CANSIM é constituída por 24 áreas temáticas principais, sendo cada uma subdividida entre 4 a 10 temas secundários. Conforme explicado na descrição deste serviço, todas as tabelas existentes são acompanhadas de um sumário que informa o possível cliente de todas as características que as constituem, sendo uma vantagem clara deste serviço. Ao escolher a tabela pretendida são ainda facultados ao utilizador ligações para artigos e outras tabelas relacionados com o tema, assim como para o catálogo do Instituto onde poderão ser adquiridos produtos sobre o tema.

StatLine é constituída por um conjunto de 7 grupos principais subdivididos em subtemas, que por sua vez ainda são divididos em grupos mais específicos. Pela

interactividade da pesquisa deste serviço, o simples passar com o “rato” por cima dos temas (ou subtemas) implica a visualização de um sumário relativo a esse tema (ou subtema). Ao ser gerada a tabela pretendida, ainda é possível clicar no seu título e obter uma janela com os seguintes conteúdos: informação geral, ligações para artigos e tabelas relacionadas (caso existam), definições, explicação de símbolos e descrição do inquérito que originou os resultados.

O serviço INFOLINE do INE possui 10 grandes temas de informação estatística que se subdividem em vários subtemas, consoante a pesquisa for feita por quadros estatísticos ou indicadores estatísticos. Normalmente, as características sobre o conteúdo dos quadros surgem com os mesmos. No que se refere a ligações para outras fontes de informação, surge no ecrã um símbolo, chamado “saber mais”, que dá acesso a uma janela com ligações para saber mais ou saber as últimas novidades estatísticas sobre o tema. Contudo esta opção apenas surge na pesquisa por indicadores estatísticos. Na pesquisa por séries cronológicas surge uma outra opção, a ficha técnica, onde o utilizador fica a conhecer todas as características da série gerada.

No critério em análise, as bases de dados CANSIM, StatLine e INFOLINE apresentam uma grande variedade de temas e, dependendo da forma de pesquisa, que poderá ser mais ou menos intuitiva ou interactiva, são apresentadas as características sobre a informação estatística produzida e ligações para outras fontes de informação relacionadas com o tema.

☒ Que quantidade de informação é fornecida em cada registo da base de dados? A quantidade de informação é útil? A quantidade de informação é suficiente para as necessidades do consumidor?

A avaliação deste critério, em cada uma das bases de dados analisadas, é difícil na medida em que é necessário estar na pele de um consumidor. Contudo, assumindo em parte este papel e ao aceder a cada uma das bases de dados, verifica-se a existência de um grande leque de informação estatística sobre os mais variados temas.

Na ausência da informação que se procura, a base de dados CANSIM permite a possibilidade de encomendar essa informação. Os serviços profissionais de CANSIM criam pacotes de dados “desenhados” de acordo com as necessidades do utilizador. Estes pacotes podem ser constituídos por tabelas, gráficos, mapas ou relatórios de análises.

A base de dados StatLine possui um método de pesquisa diferente das outras bases de dados analisadas, conforme descrito acima. É um método onde o utilizador cria a sua própria tabela (tabela dinâmica) e, por isso, é difícil não encontrar a informação

estatística desejada. Apesar de ser disponibilizada *online* toda a informação passível de ser cruzada, sob as mais variadas possibilidades, o Instituto holandês permite que o consumidor faça ainda o pedido de informação que deseja, podendo ser ou não cobrada.

No caso do serviço INFOLINE, podem ser atendidas, salvaguardando sempre o segredo estatístico, necessidades específicas de informação, inclusive por maiores níveis de desagregação e ventilação relativamente àquela que está publicada.

As bases de dados CANSIM e INFOLINE, por não possuírem um sistema baseado em quadros dinâmicos, e para irem de encontro às necessidades dos consumidores, têm um serviço de encomenda de informação estatística. Este aspecto torna-se uma vantagem em relação à StatLine, apesar desta produzir e satisfazer, *online*, o pedido do consumidor. Contudo não é uma desvantagem para a StatLine pelo facto de esta ser inovadora na forma em que chega aos utilizadores e de possuir (também) o serviço de encomenda.

☒ A base de dados é actual e está bem mantida? Com que frequência a base de dados é actualizada? Há algum atraso temporal entre a publicação da informação e a sua apresentação na base de dados?

A actualização de uma base de dados é essencial porque os utilizadores pretendem, na sua maioria, aceder às últimas novidades sobre um determinado tema e não a informações desactualizadas.

No que se refere à actualização, CANSIM possui um serviço chamado “CANSIM *New Flash*” que apresenta uma listagem diária de tabelas actualizadas. Neste serviço podem visualizar-se as actualizações por dia de semana e quais as tabelas que vão ser actualizadas no dia seguinte. Inclui ainda uma ligação para as respectivas actualizações, o que permite concluir que não possui atraso temporal na publicação da informação estatística actualizada.

A base de dados StatLine possui uma actualização praticamente diária. No *website* do Instituto é apresentada uma lista de actualizações por datas, feitas na base de dados. Assim que é feita esta divulgação, a base de dados é actualizada automaticamente, visto que, como na CANSIM, as actualizações possuem ligação directa para a tabela.

O serviço INFOLINE do INE possui uma ligação para uma página intitulada “Novidades”, onde é apresentada uma listagem de novidades por indicadores estatísticos gerais, estudos, destaques e novas publicações disponíveis. Nesta página é disponibilizada, para além da data de actualização, a ligação para a respectiva tabela, ou estudo, ou publicação que acabou de ser actualizada.

De forma a “prender” os utilizadores, os serviços analisados actualizam de uma forma sistemática e eficiente as suas bases de dados com a respectiva apresentação dos materiais para visualização e importação, caso seja pretendido.

☒ A informação está bem apresentada e organizada? Que meios de pesquisa e análise estão disponíveis? Os meios técnicos de pesquisa e de análise são úteis, efectivos e fáceis de usar? Que opções de *output* e de *download* estão disponíveis? É fácil de fornecer e descarregar dados da base de dados? Os dados podem ser exportados para outro programa? Isso é fácil de fazer?

A apresentação e organização de uma base de dados vão influenciar a facilidade com que o utilizador acede à informação. As facilidades de pesquisa constituem um factor importante na avaliação de uma base de dados em termos de eficiência e valor, assim como a forma com que é visualizada e importada a informação estatística.

A base de dados CANSIM apresenta cinco opções de pesquisa conforme referido na descrição do serviço. A pesquisa por tema apresenta a informação bem organizada e estruturada na forma de uma árvore, o que facilita a pesquisa. O *output* fornecido é constituído por tabelas, séries cronológicas ou gráficos. A possibilidade de efectuar *download* é dada em vários formatos dos quais se destacam os principais: o csv e o prn⁵².

StatLine possui a função “*Webselector*” que a distingue fortemente das outras bases de dados. Esta forma de pesquisa constitui um ponto forte e é uma forma inovadora, eficiente e extremamente fácil de utilizar. Através desta função o utilizador pode ter acesso a todos os cruzamentos de temas com subtemas que pretenda criar. As formas de visualização são tabelas, gráficos e mapas com algumas restrições no tamanho. Esta restrição, ao nível do número de células, não constitui uma desvantagem, uma vez que o limite máximo é difícil de atingir – 10 000 células *online* ou, no caso de *download*, 100 000 células. Os formatos disponíveis para *download* são vários: ficheiro *Microsoft Excel*, ficheiro *SPSS*, ficheiro de texto ou ficheiro *html*.

O INFOLINE possui algumas formas de pesquisa: por palavra-chave, por tema estatístico, ou por tipo de informação estatística. Ao optar pela pesquisa por um tipo de informação estatística, é apresentado ao utilizador uma forma de pesquisa diferente e adequada a cada tipo. As formas de visualização e importação variam consoante o tipo de informação estatística: ficheiros *html*, ficheiros *Microsoft Excel* ou ficheiros pdf.

Neste critério, a base de dados do Instituto holandês apresenta vantagens evidentes em relação às restantes abordadas. StatLine, para além de apresentar uma

⁵² PRN - Tab-separated values.

forma de pesquisa inovadora, possibilita formas de visualização interactivas e com mais um formato possível de *download*.

☒ A base de dados é facilmente acessível? Há algumas restrições ao seu acesso? Há alguma taxa para aceder à base de dados ou para descarregar registos?

CANSIM, base de dados do Instituto canadiano, é constituída por informação estatística que pode ser consultada, impressa ou importada mediante pagamento. Contudo, CANSIM possui alguma informação estatística sem custo, fornecendo ao utilizador ligações para visualizar a informação gratuita relacionada com o tema que está a pesquisar.

A base de dados StatLine possui informação estatística sobre vários temas sociais e económicos na forma de tabelas ou gráficos. Toda a informação estatística gerada pode ser consultada, impressa e importada sem qualquer custo.

No caso português, toda a informação disponível no INFOLINE é de acesso gratuito, sendo apenas necessário efectuar um registo de utilizador para aceder a determinados serviços (quadros estatísticos, séries cronológicas, pesquisa por unidade territorial, retratos territoriais e estudos). O registo pode ser efectuado através de formulário electrónico.

No que se refere a este critério, a base de dados StatLine apresenta vantagens, uma vez que disponibiliza toda a informação gratuitamente e sem exigência de registo prévio.

☒ Que impressão geral sobre a qualidade da base de dados é criada após a análise?

A avaliação e comparação de bases de dados devem considerar uma impressão geral provocada pela utilização das diferentes formas de pesquisa em termos da facilidade de uso, tendo em vista a investigação dos critérios referidos anteriormente (Cooke, 1999). Os factores que afectam a facilidade de uso da base de dados, estão intrinsecamente relacionados com a sua acessibilidade, assim como a apresentação e organização da informação. As bases de dados devem ser de fácil acesso, devendo ser fácil a movimentação dentro da mesma, assim como a localização das áreas temáticas e disponibilizarem funções de pesquisa que facilitem o seu uso. Especificamente, a sua utilização deve ser intuitiva e amigável, caso contrário, será necessário o treino para a sua familiarização antes que possa ser utilizado com eficiência. Além disto, os serviços de apoio ao utilizador devem facilitar o uso de materiais disponíveis. Como exemplos existem os guiões de utilizador ou as informações de ajuda (*help*). Tendo em consideração os conceitos referidos, a Tabela 7 resume os pontos fracos e fortes de cada uma das bases de dados analisadas.

Critérios	StatLine	CANSIM	INFOLINE
Apoio à utilização	↗	↗	↘
Conteúdo/Cobertura	↗	↗	↗
Informação sobre o <i>output</i> e Ligações a outras fontes	↗	↗	↗
Quantidade Suficiente para as Necessidades do Consumidor	↗	↗	↗
Actualização de Conteúdos	↗	↗	↗
Formas de Pesquisa	↗	↘	↘
Opções de <i>Output</i> e <i>Download</i>	↗	↘	↘
Apresentação e Organização da Informação	↗	↗	↘
Acesso à base de dados (Taxa)	↗	↘	↘↗

Tabela 7 – Pontos fortes (↗) e fracos (↘) das bases de dados StatLine, CANSIM e INFOLINE segundo alguns critérios (Cooke, 1999)

Capítulo V – Novas Perspectivas na Apresentação/Difusão de Estatísticas Oficiais

V.1 Introdução

Os métodos clássicos de análise estatística de dados foram concebidos para serem aplicados em situações relativamente simples: os dados eram obtidos para indivíduos singulares usando entrevistas, experiências, arquivos, etc.; as variáveis eram previamente definidas e apresentavam apenas um valor numérico ou uma categoria para cada indivíduo. Contudo, no mundo em constante mutação existem situações muito complexas para serem descritas apenas por estes modelos simples.

A necessidade de descrever conjuntos complexos de dados e de obter informação precisa, aliada ao objectivo de criar métodos para sumariar os dados contidos em grandes bases de dados, conduziu ao desenvolvimento da análise de dados simbólicos (Bock & Diday, 2000). Actualmente, a análise de dados simbólicos constitui uma nova perspectiva na difusão e análise de informação estatística oficial. Esta nova filosofia de análise multivariada de dados consiste numa metodologia inovadora, em contínuo desenvolvimento e aperfeiçoamento. As primeiras investigações no campo da análise de dados simbólicos tiveram início em 1988 com a publicação de um artigo de Edwin Diday (Marcelo *et al.*, 2001). Contudo, a divulgação destas técnicas de análise tiveram o seu desenvolvimento com a publicação do livro "Analysis of Symbolic Data: Exploratory Methods for Extracting Statistical Information from Complex Data" (Bock & Diday, 2000). Este surge como o relatório científico final de um projecto europeu designado de SODAS ("Symbolic Official Data Analysis System"), financiado pela Comunidade Europeia através do EUROSTAT – programa DOSIS e desenvolvido de 1996 a 1999. O propósito deste projecto foi o desenvolvimento de um *software* para análise multivariada de dados, de fácil utilização, permitindo analisar conjuntos de informação com estruturas complexas.

Após a finalização do projecto SODAS, surge o projecto ASSO⁵³, "Analysis System of Symbolic Official Data", cujo objectivo era dar continuação ao projecto anterior. O ASSO, financiado pelo EUROSTAT, no âmbito do Programa de Tecnologias da Sociedade da Informação, teve início em 2001 e uma duração de três anos. O objectivo deste projecto foi o de utilizar a análise de dados simbólicos na resolução de problemas

⁵³ ASSO - *Analysis System of Symbolic Official Data*: <http://www.assoproject.be> (consultado 10 Mar. 2005)

com que os Institutos Nacionais de Estatística se deparam, em particular problemas de confidencialidade.

Para descrever dados com estruturas complexas, como atrás referido, foram definidos novos tipos de variáveis – variáveis simbólicas – que permitem representar dados agregados, tendo em consideração a sua variabilidade e incerteza, e definir relações lógicas entre os respectivos valores. A análise de dados simbólicos permite considerar classes de objectos, descrevendo-as através das variáveis simbólicas, e analisar a tabela simbólica gerada, usando técnicas de análise adequadas.

Um dos objectivos mais importantes dos Institutos de Estatística consiste em resumir grandes conjuntos de dados em conjuntos mais pequenos, com novas unidades estatísticas, limitando o mais possível a perda de informação (Garrido, 2000). Um exemplo de uma agregação de dados é dado pela análise simbólica dos dados obtidos através do inquérito sobre a ocupação do tempo no país Basco, estudo efectuado pelo EUSTAT, Instituto Basco de Estatística (Mas & Olaeta, 2002). Os dados deste estudo referem-se ao inquérito realizado no País Basco em 1998, a 5040 indivíduos com 47 variáveis, que constituem uma lista de actividades praticadas pelos indivíduos ao longo do dia, divididas em várias áreas – emprego e estudos, família, vida social, viagens, etc. e outras que descrevem os indivíduos a nível sócio-demográfico (sexo, idade, actividade profissional). O estudo foi implementado em três momentos semanais (dias de trabalho, sextas-feiras e fins de semana) e em duas fases no ano (Primavera e Outono) de forma a abordar todas as actividades possíveis de ocupação do tempo. De forma a reduzir e a extrair informação, os autores deste estudo criaram grupos de indivíduos (cujas descrições se designam por “objectos simbólicos”), por exemplo, um constituído por indivíduos do sexo masculino com idades compreendidas entre 35 e 59 anos e trabalhadores activos. A partir desta construção e recorrendo à análise de dados simbólicos foi possível descrever e representar o comportamento dos grupos de indivíduos em cada uma das actividades de ocupação do tempo consideradas, de uma forma global.

Neste capítulo pretende-se abordar esta nova perspectiva de apresentar e difundir informação estatística, começando por definir o conceito de dado simbólico, e também perceber quais as vantagens que a aplicação deste tipo de análise traz para os Institutos Nacionais de Estatística.

V.2 Dados Simbólicos

As tabelas de dados simbólicos constituem o *input* principal da análise de dados simbólicos. Estas tabelas apresentam em coluna as variáveis simbólicas, usadas normalmente para descrever um conjunto de unidades (chamadas indivíduos), e em linha as descrições simbólicas dos indivíduos (Bock & Diday, 2000). Os dados simbólicos são dados que ultrapassam o conceito do modelo da tabela clássica, onde n indivíduos w_i , $i=1,\dots,n$ assumem apenas um valor para cada uma das p variáveis, Y_1,\dots,Y_p (Tabela 8). Agregando os indivíduos em grupos, as variáveis assumirão mais de que um único valor em cada grupo, surgindo assim os dados simbólicos que preenchem as células da referida tabela simbólica.

Variáveis sócio-demográficas				Variáveis Ocupação do Tempo				Ponderação Amostral
ID	Sexo	Idade	...	Dia	Limpar	Cozinhar	Dormir	
1	1	3	...	2	SP	4	2	...
2	2	4	...	1	PP	3	2	...
...

Tabela 8 – Tabela do inquérito sobre a ocupação do tempo no país Basco

(Fonte: Mas & Olaeta, 2002)

Para representar dados simbólicos, novos tipos de variáveis foram definidos: variáveis multi-valores, variáveis intervalares e variáveis modais (Bock & Diday, 2000). Uma variável Y é chamada de *set-valued* se os seus valores são conjuntos não vazios do seu domínio, de multi-valores se os seus valores forem subconjuntos finitos do seu domínio (valores qualitativos ou quantitativos) e intervalar se os seus valores tomarem a forma de intervalos. Uma variável modal Y_i com domínio finito $O_j=\{m_1,\dots,m_{k_j}\}$ é uma variável multi-condição que a cada elemento faz corresponder um conjunto de categorias e para cada categoria m_i é indicada a frequência ou probabilidade p_i , que indica o quão frequente é a categoria para o respectivo elemento (Brito, 2002).

Estes tipos de variáveis serão ilustrados, em seguida, com exemplos.

Exemplo de variáveis multi-valores

Num estudo europeu sobre a distribuição de instituições bancárias pelas cidades europeias, pode definir-se a variável $Y(k)$ ="instituições bancárias presentes na cidade k " cujo domínio Y é constituído pelas partes do conjunto das possíveis instituições bancárias, $Y=\{Barc, BNP, CL, Spk, DB, Dres, BdiRoma, Lux,\dots\}$. O resultado poderá ser a tabela simbólica Tabela 9.

K (cidade)	Y(k)
Paris	{BNP, CL, DB, Dres}
Bona	{Spk, DB, Dres, Lux}
Londres	{Barc, CL, DB}
Roma	{DB, BdiRoma}
Luxemburgo	Y

Tabela 9 – Tabela Simbólica para uma variável multi-valores (Fonte: Bock & Diday, 2000)

Exemplo de variáveis intervalares

Num estudo sobre a educação, pode-se considerar o conjunto E formado pelos estudantes que participam num determinado curso e a variável Y que representa o tempo despendido pelos estudantes na aprendizagem (em horas). As respostas possíveis podem ser representadas na forma de intervalo, por exemplo: $Y(k)=[0, 2]$, $Y(l)=[.5, 4.6]$.

Exemplo de variáveis modais

Num inquérito sobre condições de emprego, considere-se a variável "nível de instrução". Para descrever um grupo k particular (por exemplo, mulheres de 30 a 40 anos), podem indicar-se as frequências dos vários níveis de ensino; os valores da variável modal Y resultante são distribuições de frequências:

$Y(k) = \{\text{primário (30\%), secundário (90\%), superior (10\%)}\}$.

Todos estes tipos de variáveis, em conjunto com as variáveis clássicas (que assumem valores singulares) permitem representar aquilo a que se chamam dados simbólicos. A determinação, observação ou registo de dados conduz à utilização de uma matriz de dados simbólicos. Assim, os valores que constituem uma matriz de dados simbólicos não são, em geral, valores singulares mas sim conjuntos, intervalos, distribuições de frequências, etc. Como exemplo, considere-se um conjunto E constituído por 4 cidades, a_1, a_2, a_3, a_4 e as seguintes variáveis simbólicas Y_1, Y_2, Y_3 :

Y_1 : =número de habitantes (mínimo e máximo relativos aos anos 1990-1995, em milhares) com valores na forma de intervalo $[a, b] = [\min, \max]$

Y_2 : =partidos políticos numa cidade (DC=Democrata-Cristão, SD=Social-Democrata, S=Socialista e C=Comunista) com valores percentuais obtidos na última eleição.

Y_3 : =lista das grandes instituições bancárias existentes na cidade (subconjunto da lista total de instituições bancárias {BCP, BES, TOTTA, BIC}).

Com estas 3 variáveis é possível construir uma matriz de dados simbólicos (4x3) onde cada linha descreve uma cidade:

$$X = \begin{pmatrix} (80, 100] & (DC 0.2; SD 0.3; S 0.4; C 0.1) & \{BCP, BIC\} \\ (100, 130] & (DC 0.1; SD 0.3; S 0.4; C 0.2) & \{BCP, BES, TOTTA\} \\ (8, 10] & (DC 0.1; SD 0.5; S 0.3; C 0.1) & \{BES\} \\ (10, 13] & (DC 0.1; SD 0.3; S 0.5; N 0.1) & \{BCP, TOTTA, BIC\} \end{pmatrix}.$$

V.3 Análise de Dados Simbólicos nos Institutos Nacionais de Estatística

A necessidade de retirar informação a partir de dados cada vez mais complexos tem vindo a crescer de uma forma gradual. Assim, a partir do momento em que se possui uma grande quantidade de informação é importante agregá-la de forma a construir conceitos. A análise simbólica de dados assume aspectos que a tornam uma boa opção para ser aplicada nas estatísticas oficiais pois fornece a capacidade de descrever dados com uma estrutura complexa. A informação é extraída da base de dados e será descrita através de objectos simbólicos que descrevem indivíduos ou classes de indivíduos. Em comparação com a abordagem clássica, os métodos da análise de dados simbólicos apresentam vantagens quando aplicados em Institutos Nacionais de Estatística. Os principais campos de aplicação desta análise nos Institutos Nacionais de Estatística são o controlo de qualidade durante o desenvolvimento de um estudo e a exploração das bases de dados finais (Marcelo & Brito, 2004). Para além destes métodos poderem ser aplicados a dados mais complexos, permitem também ultrapassar o problema da confidencialidade de dados pessoais, utilizar intervalos de confiança e possibilitam a junção de bases de dados independentes.

Conforme referido no Capítulo IV, os Institutos Nacionais de Estatística não podem por lei divulgar informação individual, o que os obriga a agregar os dados antes de os difundir a outras instituições ou ao público em geral. Como a análise de dados simbólicos permite gerir a agregação de dados, os institutos poderão proceder previamente a uma agregação para fornecer dados aos investigadores externos sem problemas de confidencialidade. Por outro lado, uma tabela simbólica contém mais

informação sobre a variabilidade das variáveis do que uma simples tabela de contingência, usada frequentemente pelos institutos (Marcelo & Brito, 2004) – ver Figura 39.

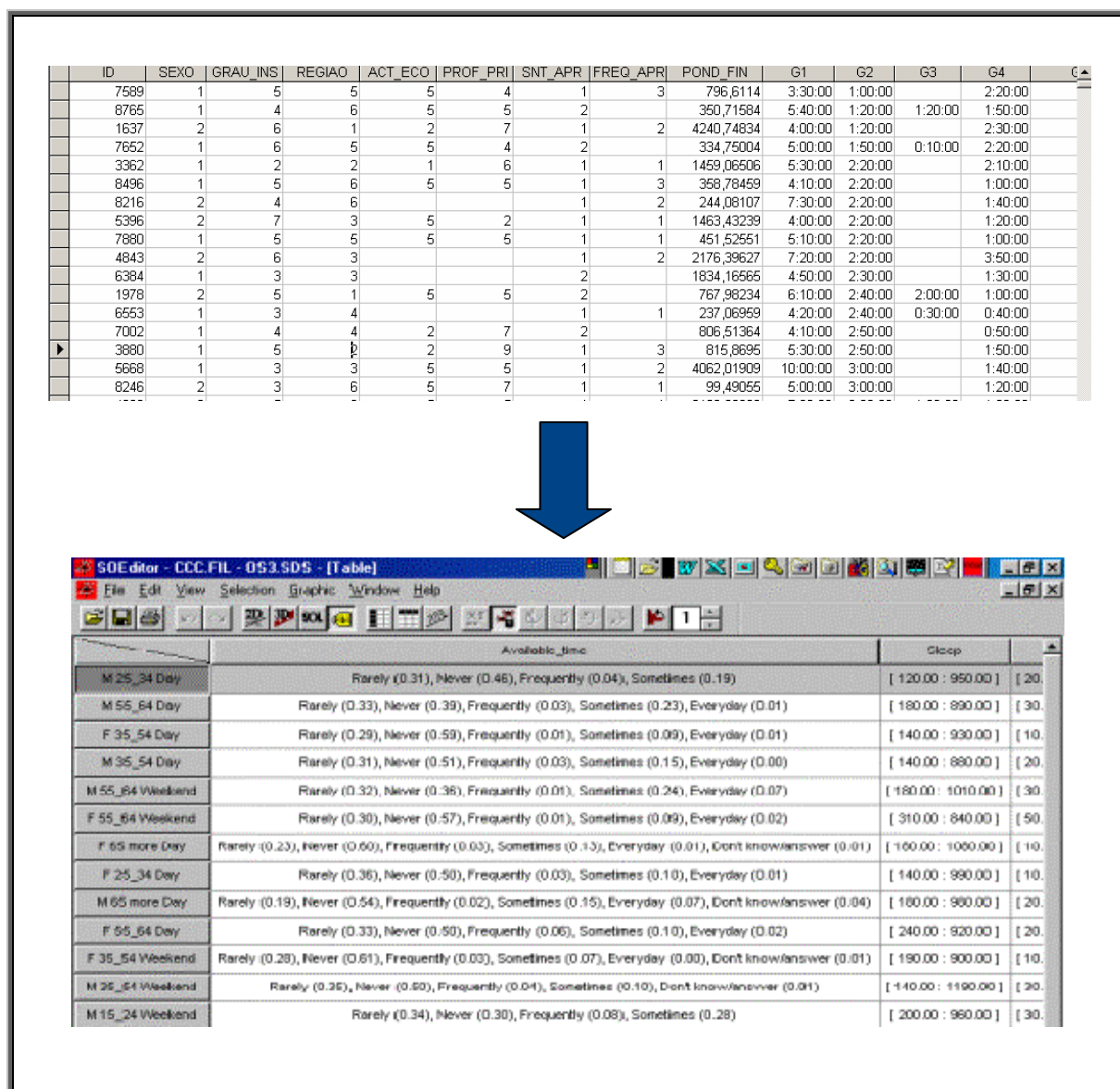


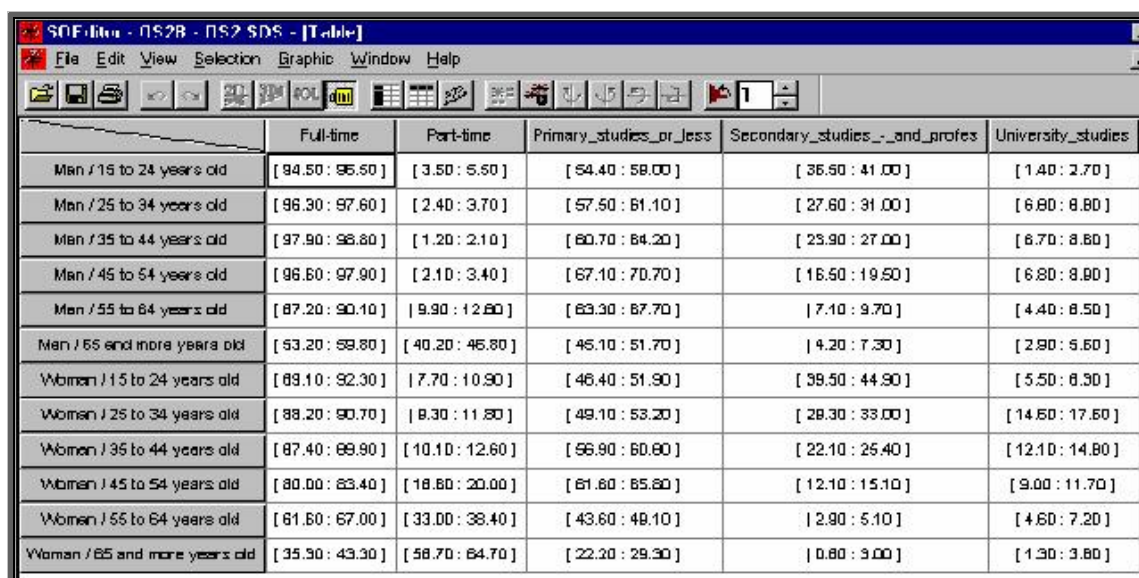
Figura 39 – Exemplo de microdados e respectivos dados agregados

Os Institutos Nacionais de Estatística recolhem a maioria dos dados através de estudos amostrais, ou seja, apenas uma parte da população é entrevistada. Neste contexto, as análises feitas com base em intervalos de confiança constituem um aspecto importante para obter a informação desejada. Na análise clássica apenas é possível colocar um único valor em cada célula da tabela, enquanto que na análise de dados simbólicos é possível trabalhar com dados intervalares (Marcelo & Brito, 2004). Como

exemplo prático desta vantagem, refere-se o inquérito sobre a actividade laboral realizado em 1998 pelo INE (Marcelo, 2004), a 22660 trabalhadores. Neste estudo foram recolhidas variáveis sócio-demográficas (grupo etário, sexo, estado civil, nível de escolaridade) e variáveis relacionadas com a actividade laboral (actividade económica, profissão, procura de emprego, tempo inteiro ou parcial, tipo de empresa, etc.). As novas unidades construídas (por exemplo) com base nas variáveis Sexo (2 categorias) e grupo etário (6 categorias), representam 2x6 classes de indivíduos. Cada uma destas classes, representada por um objecto simbólico, é constituída por indivíduos de um dado sexo e grupo etário, por exemplo, indivíduos do sexo masculino com idades entre os 15 e 24 anos. Para cada uma das modalidades (j) das variáveis relacionadas com a actividade laboral e para cada objecto simbólico (i) é construído um intervalo de confiança a 95% cujos extremos, a_{ij} e b_{ij} são calculados da seguinte forma:

$$a_{ij} = p_{ij} - 1.96 * \sqrt{\frac{p_{ij}(1-p_{ij})}{n_i}} \quad b_{ij} = p_{ij} + 1.96 * \sqrt{\frac{p_{ij}(1-p_{ij})}{n_i}}, \text{ onde } n_i \text{ representa o número}$$

de indivíduos que constituem a classe representada pelo objecto simbólico i e p_{ij} a proporção de indivíduos com a categoria j nesta classe. Como resultado da agregação de dados é construída a tabela simbólica representada na Figura 40. A partir da tabela simbólica pode-se representar graficamente a informação recorrendo a um *software* adequado, abordado no ponto 4 deste capítulo, e daí retirar informações pertinentes.



	Full-time	Part-time	Primary_studies_or_less	Secondary_studies_and_profes	University_studies
Men / 15 to 24 years old	[94.50 : 95.50]	[3.50 : 5.50]	[54.40 : 59.00]	[35.50 : 41.00]	[1.40 : 2.70]
Men / 25 to 34 years old	[96.30 : 97.60]	[2.40 : 3.70]	[57.50 : 61.10]	[27.60 : 31.00]	[6.60 : 8.80]
Men / 35 to 44 years old	[97.90 : 98.80]	[1.20 : 2.10]	[60.70 : 64.20]	[25.90 : 27.00]	[6.70 : 8.80]
Men / 45 to 54 years old	[96.60 : 97.90]	[2.10 : 3.40]	[67.10 : 70.70]	[16.50 : 19.50]	[6.80 : 8.80]
Men / 55 to 64 years old	[87.20 : 90.10]	[9.90 : 12.80]	[63.30 : 67.70]	[7.10 : 9.70]	[4.40 : 6.50]
Men / 65 and more years old	[53.20 : 59.80]	[40.20 : 45.80]	[45.10 : 51.70]	[4.20 : 7.30]	[2.90 : 5.60]
Women / 15 to 24 years old	[89.10 : 92.30]	[7.70 : 10.90]	[46.40 : 51.90]	[39.50 : 44.90]	[5.50 : 6.30]
Women / 25 to 34 years old	[88.20 : 90.70]	[8.30 : 11.80]	[49.10 : 53.20]	[28.30 : 33.00]	[14.60 : 17.60]
Women / 35 to 44 years old	[87.40 : 89.90]	[10.10 : 12.60]	[56.90 : 60.80]	[22.10 : 25.40]	[12.10 : 14.80]
Women / 45 to 54 years old	[80.00 : 83.40]	[18.60 : 20.00]	[61.60 : 65.60]	[12.10 : 15.10]	[9.00 : 11.70]
Women / 55 to 64 years old	[61.60 : 67.00]	[33.00 : 38.40]	[43.60 : 49.10]	[2.90 : 5.10]	[4.60 : 7.20]
Women / 65 and more years old	[35.30 : 43.30]	[58.70 : 64.70]	[22.20 : 29.30]	[0.80 : 3.00]	[1.30 : 3.80]

Figura 40 – Tabela simbólica com intervalos de confiança (Fonte: Marcelo, 2004)

A possibilidade de juntar estudos independentes constitui também uma vantagem para os Institutos Nacionais de Estatística. Frequentemente, surge o interesse de desenvolver análises baseadas em variáveis oriundas de dois ou mais estudos. Para efectuar este tipo de análise é necessário ter uma tabela em que as linhas representam os indivíduos e as colunas as variáveis. Contudo, quando os indivíduos não são os mesmos em ambos os estudos, os microdados não podem ser agregados. Na análise de dados simbólicos, as unidades são classes de indivíduos, agrupados de acordo com um determinado critério. Se as mesmas classes de indivíduos forem definidas em todas as bases de dados, de acordo com a análise pretendida, então a descrição dessas classes, relativa a diferentes estudos, pode ser agregada numa única tabela simbólica. Por exemplo, um estudo sobre as condições de vida e outro sobre a ocupação do tempo podem ser juntos formando uma única tabela simbólica (Marcelo, 2004), desde que se definam os mesmos grupos. Supondo que em ambos os estudos se pretende perceber o comportamento das mulheres casadas com 35 a 59 anos de idade, com formação universitária e sendo trabalhadoras activas, a análise de dados simbólicos permite extrair a informação sob a forma de *Zoom-Star*, representada na Figura 41.

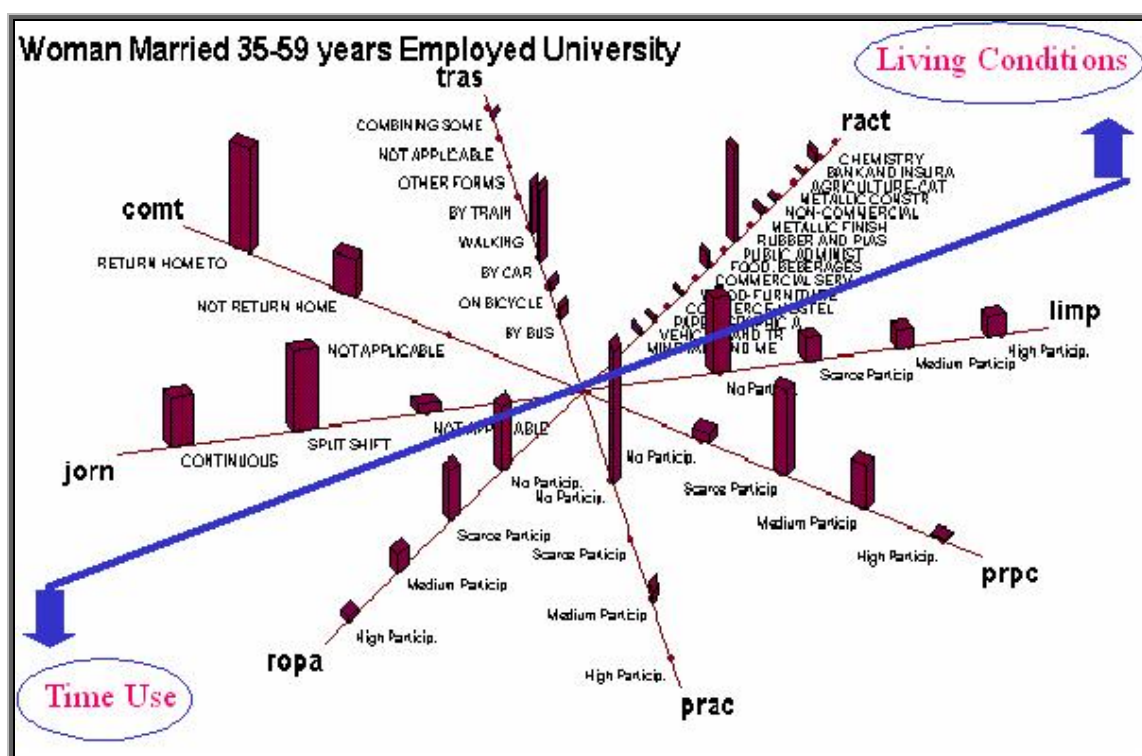


Figura 41 – Exemplo de junção de dois estudos independentes através da análise de dados simbólicos; representação por *zoom-star* (Fonte: Marcelo, 2004)

A análise de dados simbólicos parece ainda ser uma poderosa ferramenta para desenvolver aplicações no processo de controlo de qualidade da informação estatística produzida nos Institutos Nacionais de Estatística. A produção de informação estatística, através de estudos amostrais, engloba várias fases e intervenientes, até à construção da base de dados final: a metodologia (método de amostragem, determinação da dimensão da amostra); o desenho do questionário; o processo de recolha; a introdução, codificação e validação dos dados e o cálculo de estimadores e de erros de amostragem. Assim, o desenvolvimento de um estudo estatístico é um processo complexo que pode gerar erros, especialmente os não relacionados com a amostragem. A implementação de um sistema de controlo associado ao processo de produção de informação estatística, baseado no cálculo de indicadores para cada uma das etapas do estudo estatístico, é crucial para perceber quais as áreas que necessitam de ser melhoradas, de forma a minimizar os erros. Juntamente com o cálculo destes indicadores, a análise de dados simbólicos pode ser aplicada na análise da *reentrevista* (para avaliar os erros cometidos na recolha e digitação dos dados - erros de medida), na análise de *recodificação* (para avaliar a codificação – erros de processamento) e na validação da amostra através de rotação (Marcelo & Brito, 2004) (Figura 42).

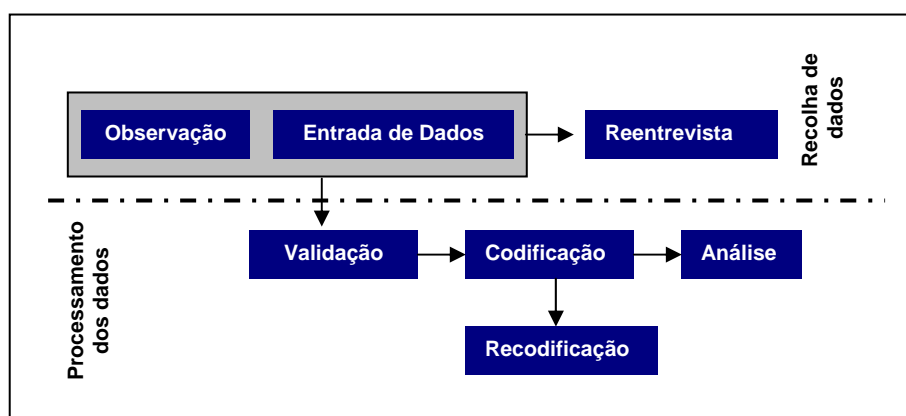


Figura 42 – Processo de controlo de qualidade na produção de informação estatística

(Fonte: Marcelo, 2004)

Considere-se o Inquérito ao Emprego da zona de Lisboa e Vale do Tejo, Portugal, para o 1º e 2º trimestre de 2000. Do vasto conjunto de informação passível de extrair deste inquérito, restringiu-se a análise às variáveis que permitem caracterizar os empregados no seu conjunto (sócio-demográficas: sexo, grupo etário, estado civil, nível de escolaridade; actividade laboral: ocupação, actividade económica, categoria profissional, tempo inteiro ou parcial, número de horas no grupo de trabalho), focalizando-

se nos 8669 indivíduos presentes na amostra deste inquérito que se encontravam empregados. Considerando a rotação da amostra – 6 partes de um trimestre x 2 trimestres – o número de objectos simbólicos a criar será de 6x2; a rotação consiste em eliminar a primeira parte do primeiro trimestre, substituindo-a por uma nova sexta parte no trimestre seguinte (Figura 43). Representando cada parte de cada trimestre por um objecto simbólico, é possível fazer a sua descrição numa tabela simbólica e, posteriormente aplicar a análise com vista a controlar a estabilidade da amostra.

1	2	3	4	5	6	1º Trimestre
1/1	1/2	1/3	1/4	1/5	1/6	
	2	3	4	5	6	7
	2/2	2/3	2/4	2/5	2/6	2/7

Figura 43 – Processo de rotação da amostra (Fonte: Marcelo, 2004)

rot	sex	age_group	marital_status	economic_activity
1 / 2	Male	55 <= age <= 64	Married	Wholesale and retail trade, re
1 / 6	Male	35 <= age <= 44	Single	Construction
1 / 1	Male	35 <= age <= 44	Married	Manufacturing
1 / 2	Male	35 <= age <= 44	Married	Manufacturing
1 / 2	Female	25 <= age <= 34	Married	Other Services
1 / 2	Male	45 <= age <= 54	Married	Manufacturing
1 / 2	Female	45 <= age <= 54	Married	Manufacturing
1 / 2	Male	15 <= age <= 24	Single	Manufacturing
1 / 2	Female	15 <= age <= 24	Single	Manufacturing
1 / 2	Female	15 <= age <= 24	Single	Manufacturing
1 / 6	Male	35 <= age <= 44	Married	Manufacturing
1 / 6	Female	35 <= age <= 44	Married	Manufacturing

	sex	age_group
2 / 6	Female (0.45), Male (0.55)	45 <= age <= 54 (0.22), 35 <= age <= 44 (0.28), 55 <= age <= 64 (0.08), age >= 65 (0.02), 25 <
2 / 5	Female (0.46), Male (0.54)	45 <= age <= 54 (0.22), 35 <= age <= 44 (0.25), 55 <= age <= 64 (0.14), age >= 65 (0.03), 25 <
2 / 3	Female (0.47), Male (0.53)	45 <= age <= 54 (0.25), 35 <= age <= 44 (0.22), 55 <= age <= 64 (0.12), age >= 65 (0.02), 25 <
2 / 4	Female (0.45), Male (0.55)	45 <= age <= 54 (0.22), 35 <= age <= 44 (0.22), 55 <= age <= 64 (0.14), age >= 65 (0.04), 25 <
2 / 7	Female (0.47), Male (0.53)	45 <= age <= 54 (0.23), 35 <= age <= 44 (0.24), 55 <= age <= 64 (0.12), age >= 65 (0.04), 25 <
1 / 1	Female (0.45), Male (0.55)	45 <= age <= 54 (0.22), 35 <= age <= 44 (0.26), 55 <= age <= 64 (0.13), age >= 65 (0.03), 25 <
2 / 2	Female (0.47), Male (0.53)	45 <= age <= 54 (0.23), 35 <= age <= 44 (0.24), 55 <= age <= 64 (0.13), age >= 65 (0.03), 25 <
1 / 5	Female (0.46), Male (0.54)	45 <= age <= 54 (0.21), 35 <= age <= 44 (0.25), 55 <= age <= 64 (0.14), age >= 65 (0.02), 25 <
1 / 2	Female (0.47), Male (0.53)	45 <= age <= 54 (0.23), 35 <= age <= 44 (0.24), 55 <= age <= 64 (0.13), age >= 65 (0.03), 25 <
1 / 3	Female (0.47), Male (0.53)	45 <= age <= 54 (0.25), 35 <= age <= 44 (0.21), 55 <= age <= 64 (0.12), age >= 65 (0.02), 25 <
1 / 6	Female (0.44), Male (0.56)	45 <= age <= 54 (0.21), 35 <= age <= 44 (0.28), 55 <= age <= 64 (0.08), age >= 65 (0.02), 25 <
1 / 4	Female (0.45), Male (0.55)	45 <= age <= 54 (0.23), 35 <= age <= 44 (0.21), 55 <= age <= 64 (0.13), age >= 65 (0.03), 25 <

Figura 44 – Tabela clássica e tabela simbólica em que os objectos simbólicos representam as partes da amostra na 1ª e 2ª rotações (Fonte: Marcelo, 2004)

V.4 Software SODAS

SODAS, *Symbolic Official Data Analysis System*, foi o primeiro projecto europeu cujo objectivo principal consistiu no desenvolvimento de métodos de análise estatística e *software* para análise de dados complexos definidos sob a forma de objectos simbólicos. Este projecto foi desenvolvido entre 1997 e 1999, e resultou na edição de um livro (Bock & Diday, 2000) e no *software* SODAS.

O objectivo do *software* SODAS é o de facilitar a utilização de técnicas de análise simbólica de dados nos Institutos Nacionais de Estatística e empresas, e consequentemente demonstrar que estas técnicas vão de encontro a várias necessidades dos utilizadores: analisar dados com uma estrutura complexa, obter resultados estatísticos com melhor apresentação e fácil interpretação e ainda descrever, unificar e analisar conceitos equivalentes de estatísticas oficiais de diferentes países. Do projecto ASSO⁵⁴, continuação do projecto SODAS, resultou um *software* designado SODAS2, que consiste num aperfeiçoamento da versão anterior, seguindo pedidos e sugestões de utilizadores. SODAS2 é mais operacional e atractivo e oferece métodos inovadores.

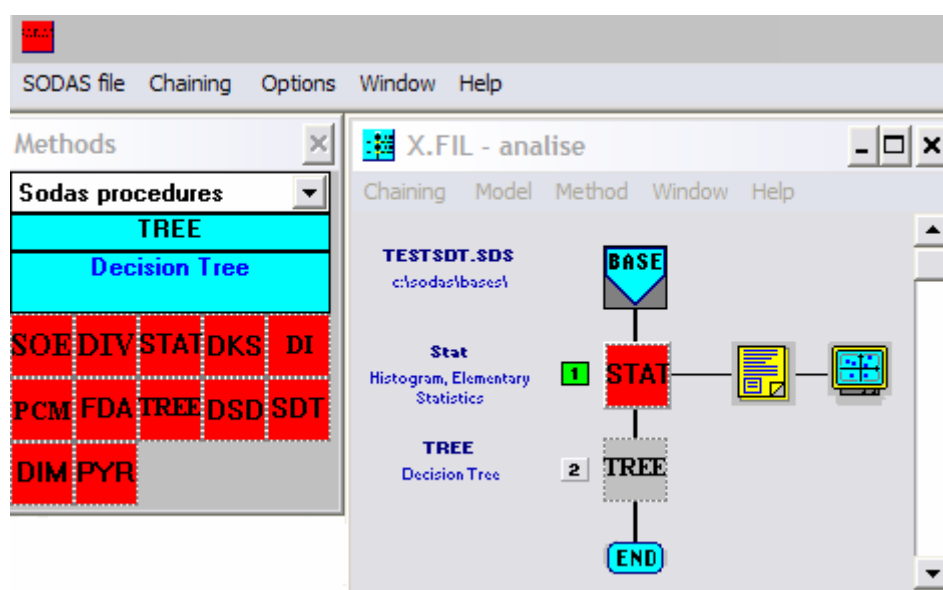


Figura 45 – Ambiente de trabalho do software SODAS, versão 1.2

Este *software* inclui ferramentas para criar, guardar e actualizar objectos simbólicos, criar objectos simbólicos a partir de outras bases de dados (DB2SO), criar

⁵⁴ ASSO - Analysis System of Symbolic Official Data: <http://www.assoproject.be> (consultado 10 Mar. 2005)

objectos simbólicos a partir de dados originais (ND2SO), exportar objectos simbólicos para bases de dados, e uma colecção de métodos de análise estatística: métodos estatísticos descritivos, análise classificatória, árvores de decisão, análise discriminante, regressão e análise factorial (Figura 46).

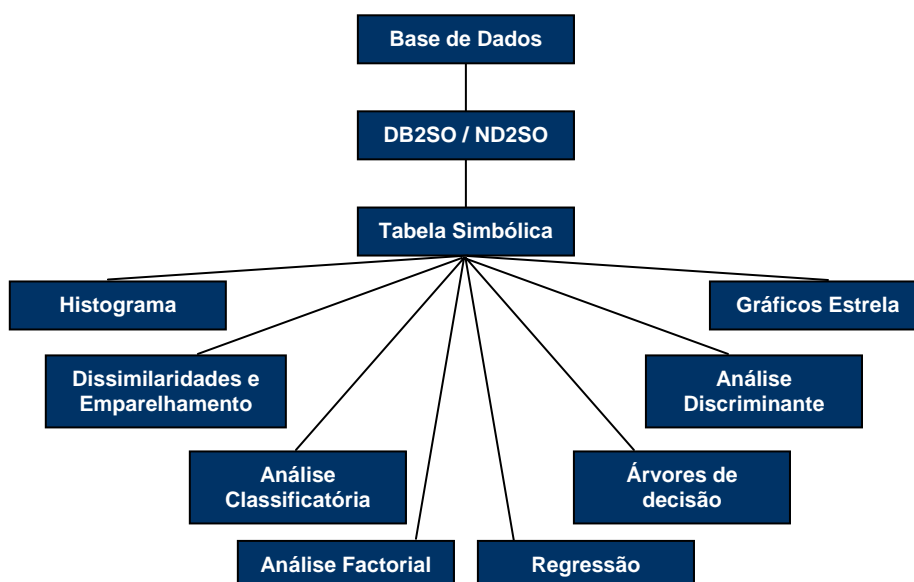


Figura 46 – Visão geral do *software* SODAS2 (Fontes: Bock & Diday, 2000; ASSO Developers, 2005)

	num_tournoi	code_surface	Classification	Dotation
Dur	[1.00 : 66.00]	DUR (1.00)	Int. (0.65), Maste (0.22), Grand (0.09), Int. (0.04)	[375000.00 : 5819900.00]
Gazon	[32.00 : 39.00]	GAZ (1.00)	Int. (0.83), Grand (0.17)	[400000.00 : 5819900.00]
Indoor	[7.00 : 65.00]	IND (1.00)	Int. (0.64), Maste (0.14), Int. (0.21)	[400000.00 : 2950000.00]
Terre-Battue	[10.00 : 56.00]	TBT (1.00)	Int. (0.70), Maste (0.13), Grand (0.04), Int. (0.13)	[400000.00 : 5819900.00]

Figura 47 – Exemplo de uma tabela simbólica no *software* SODAS. (Fonte: Exemplos SODAS)

V.5 Estudos Já Efectuados

Vários estudos foram até à data efectuados com recurso à análise de dados simbólicos e com a utilização do *software* SODAS. Em cada um deles, é efectuada a transformação dos dados que estão em tabelas clássicas, por agregação, para as tabelas simbólicas, às quais são aplicadas as técnicas da análise de dados simbólicos.

Em (Bock & Diday (2000)) são apresentados dois exemplos de aplicação da análise de dados simbólicos em estatísticas oficiais, de três institutos de estatística

envolvidos no projecto SODAS: ONS (Office for National Statistics, Inglaterra), INE (Instituto Nacional de Estatística, Portugal) e EUSTAT (Instituto de Estatística do País Basco, Espanha).

O exemplo inglês ilustra a aplicação de métodos disponíveis no *software* SODAS a dados recolhidos nos censos. Os dados estatísticos referem-se a 374 regiões do Reino Unido descritas por 34 variáveis, que abordam temas como a habitação, o agregado familiar, o emprego, etc. Os valores recolhidos foram transformados em intervalos de confiança da forma descrita no ponto 3 deste capítulo. Após ser obtida a tabela simbólica, foi aplicada uma análise classificatória descendente, de forma a obter uma árvore de decisão descrevendo algumas classes. Trata-se assim de um exemplo em que os intervalos de confiança estimados são tratados como unidades de dados simbólicos.

O segundo exemplo refere-se à comparação dos dados obtidos de dois inquéritos ao emprego, um efectuado em Portugal e outro no País Basco. Este exemplo reflecte a possibilidade de juntar estudos independentes, uma das vantagens que a análise de dados simbólicos proporciona. Os dados referem-se a 56 049 registos que descrevem indivíduos através de 17 variáveis. Estas variáveis foram adaptadas por ambos os institutos de forma a estabelecer uma coerência entre elas e para obter a mesma estrutura antes de os juntar numa tabela clássica. O objectivo consistiu em construir objectos simbólicos regionais e sócio-demográficos de forma a caracterizar regiões e sub-populações de ambos os países. A análise de dados simbólicos permite, por exemplo, comparar as distribuições das profissões, ramos de actividade económica, etc., em diferentes regiões através dos gráficos estrela. A mesma metodologia foi utilizada para comparar diferentes grupos sócio-demográficos.

Um outro caso onde foi aplicada a análise de dados simbólicos, num estudo português, foi no vasto conjunto de informação do Inquérito ao Emprego, realizado no segundo trimestre de 1998, cujo objectivo era caracterizar os mercados de trabalho ao nível das NUTS II (Nomenclatura de Unidades Territoriais para fins Estatísticos – nível II), dando ênfase à análise das características dos empregados (Marcelo, Silva & Rodrigues, 2001). Foram apresentados os retratos territoriais de emprego ao nível de NUTS II, elaborados com base na aplicação de diversos métodos de análise multivariada aos objectos simbólicos construídos. As análises efectuadas tiveram como base diferentes tabelas de dados simbólicos, consoante o caso em estudo, em que o número de linhas era o número de objectos simbólicos (dependendo da análise em causa) e as colunas as variáveis intervalares associadas (intervalos de confiança a 95% para a proporção). Por exemplo, para o caso em que o objecto simbólico é constituído apenas com base na

região, foi construída uma caracterização para cada uma das 7 regiões consideradas. Para uma primeira caracterização do emprego regional com vista à apresentação das suas especificidades, recorreu-se aos gráficos em estrela dos valores (intervalares) de cada categoria para cada objecto simbólico. Foi ainda efectuada uma análise factorial em componentes principais de forma a reduzir a dimensão da informação e caracterizar os objectos simbólicos. Para avaliar as similaridades/dissimilaridades regionais de emprego foi aplicada a análise classificatória. Segundo os autores, os resultados obtidos conduzem ao encorajamento da utilização da análise de dados simbólicos a outras áreas.

Por último refere-se ainda um estudo realizado pelo EUSTAT com o objectivo de caracterizar os utilizadores de Internet no País Basco (Mas *et al.*, 2002). Este estudo teve um percurso idêntico aos estudos anteriores no que se refere à passagem da tabela clássica para a tabela simbólica. A metodologia utilizada foi a construção de gráficos estrela que representam o perfil do utilizador e do não utilizador da Internet e do e-comprador. Os autores deste estudo concluem que a representação simbólica proporciona uma abordagem proveitosa na exploração de dados obtidos através de grandes estudos e permite visualizar facilmente as características dos objectos simbólicos criados.

V.6 Conclusão

A análise de dados simbólicos nasceu a partir da influência de vários campos de investigação: análise de dados exploratória, inteligência artificial e taxonomia numérica na área de Biologia (Bock & Diday, 2000). A evolução desta análise surgiu, fundamentalmente, com a necessidade de agrupar informação extensa numa única tabela, ou seja, de criar grupos com as respectivas descrições. Esta nova abordagem permite estender as tradicionais técnicas de análise estatística aos dados representados por variáveis simbólicas que podem assumir valores como subconjuntos de categorias, intervalos ou distribuições de frequências, ao contrário da análise clássica, onde as variáveis apenas podem assumir um único número ou categoria. A principal vantagem reside no facto de, com recurso à análise de objectos simbólicos, ser possível analisar as complexas bases de dados dos nossos dias, que raras vezes preenchem os requisitos necessários à análise clássica. Para além disso, a agregação dos dados iniciais sob a forma de objectos simbólicos, apresentados em tabelas simbólicas, permite preservar a confidencialidade dos microdados, constituindo assim uma vantagem evidente para os Institutos Nacionais de Estatística.

O projecto SODAS, no qual várias instituições de nove países europeus participaram, permitiu produzir o primeiro *software* para análise de dados simbólicos. Três institutos nacionais de estatística estiveram envolvidos neste projecto: o Instituto Nacional de Estatística do País Basco (EUSTAT), o Instituto Nacional de Estatística de Portugal (INE) e o Instituto Nacional de Estatística inglês (ONS) (Diday, 2002). A criação deste *software* para analisar dados simbólicos permite obter informação estatística através de um ambiente gráfico fácil de utilizar, para além de fornecer resultados cuja interpretação é clara.

Os exemplos abordados ao longo do capítulo permitem concluir que esta nova forma de apresentar e analisar os dados permite avançar no desenvolvimento do mundo estatístico e, assim, chegar mais perto do cidadão comum.

Capítulo VI – Estudo de um Caso: Serviço INFOLINE do INE

Parte A – Metodologia

VI.1 Introdução

O estudo de caso que se descreve neste capítulo teve como base a realização de um inquérito relativo ao serviço INFOLINE do INE.

VI.2 Inquérito

O questionário foi elaborado com os objectivos de conhecer o grau de utilização da Internet como fonte de informação estatística e, em particular, o grau de utilização do serviço INFOLINE do INE. Pretendia-se ainda verificar se esta utilização poderia contribuir para a satisfação de necessidades, a nível empresarial, de acesso à informação estatística.

Como é sabido, não há nunca uma opinião verdadeira, independentemente das condições da sua recolha. Segundo Ghiglione & Matalon (1993, pág.2), “os discursos que constituem a matéria-prima do inquérito não são espontâneos; (...) é necessário ter em conta o facto de não termos qualquer razão para admitir que a pessoa se submeterá passivamente às recomendações do inquiridor e dirá directamente a “verdade”, nem mesmo, mais modestamente, “a sua verdade”. Conscientemente ou não, ela diz-nos apenas o que *pode e quer* dizer-nos, facto que é determinado pela *representação* que faz da situação e pelos seus próprios objectivos, que não coincidem necessariamente com os do investigador.”

Logo, e independentemente dos cuidados havidos com o tipo de amostra, há margem para erro e, como tal, deve ser tida em conta. Nesse sentido, procurou-se garantir uma margem de erro na representatividade que não ultrapassasse os 7%⁵⁵. Assim, a amostra inicial é formada por duzentas grandes empresas de vários sectores de actividade e zonas do país, às quais foi enviado um exemplar de um questionário, acompanhado de um envelope selado para envio do questionário preenchido. O envio

⁵⁵ Ao estimar uma proporção com 95% de confiança, uma amostra de 200 observações fornece um intervalo com erro máximo (semi-amplitude) de aproximadamente 7%.

dos questionários realizou-se em Maio de 2004 e destes, apenas cinquenta e nove, recebidos até Agosto de 2004, constituíram a amostra final.

Não se pode deixar de salientar que a construção de uma amostra representativa⁵⁶ de uma população é algo complicado de realizar⁵⁷, sobretudo atendendo ao elevado custo que frequentemente não está ao alcance de quem pretende realizar um trabalho como este.

Na elaboração do questionário houve sempre a preocupação de garantir o anonimato, de utilizar uma linguagem clara e directa, de evitar perguntas múltiplas e indefinidas, de elaborar questões neutras e de evitar escrever perguntas com mistura de conjunções e disjunções (Hill & Hill, 2002). Houve ainda o cuidado de apresentar um “layout” claro e atraente e de estabelecer uma relação coerente entre a aparência estética e o tamanho do questionário.

Assim, com a elaboração deste questionário pretendia-se saber, ao nível de utilizadores do serviço INFOLINE do INE:

- como tomaram conhecimento do serviço, questão 3;
- qual a frequência de utilização do serviço, questão 4;
- qual o tipo e tema de informação estatística acedidos, questões 5 e 6;
- qual a apreciação do serviço, questão 7;
- se existe satisfação ao nível de procura de informação estatística e, no caso de não haver, qual o motivo, questão 8.

Ao nível de utilizadores e não utilizadores do serviço INFOLINE, pretendia-se saber:

- se acedem a outros *sites* de informação estatística, questão 9;
- se sim, qual ou quais os *sites* a que acedem, questão 9;
- se não, qual o motivo de não aceder, questão 9.

O questionário elaborado é constituído por questões abertas e questões fechadas. As questões abertas referem-se a situações cujas respostas poderiam dar mais informação e que, à partida, iriam produzir informação inesperada (Hill & Hill, 2002). O tratamento destas questões originou a necessidade de categorizar previamente as respostas.

⁵⁶ Uma amostra é representativa quando o grupo de indivíduos/empresas é escolhido de forma a permitir que as conclusões que dele sejam feitas se possam generalizar à totalidade da população. Para que seja representativa, a amostra deve “apresentar características idênticas às da população” (Ghiglione & Matalon, 1993, pág. 32).

⁵⁷ A amostra considerada e a respectiva listagem das empresas foi obtida através da colaboração do INE.

Como habitualmente acontece nos trabalhos de investigação por questionário, foi elaborado um pré-teste com o objectivo de avaliar a adequação do questionário a utilizar (Ghiglione & Matalon, 1993). Após este estudo, que envolveu a colaboração de alguns colegas, procurando verificar a familiarização com o vocabulário que se pretendia utilizar (Pardal & Correia, 1995), foram feitos alguns acertos de que resultou o questionário final, o qual é apresentado no Anexo 2.

Quanto à modalidade das questões, natureza das variáveis e escalas de atitudes utilizadas tem-se (Javeau, 1978; Sachs, 1984):

- abertas (parte das questões 8 e 9);
- múltiplas em leque aberto, de escala nominal (questões 3 e 4);
- fechada dicotómica, de escala nominal (questão 2);
- múltiplas em leque fechado, de escala nominal (questões 5 e 6);
- de avaliação, sem ponto neutro, com escala de atitudes de Likert, escala ordinal (questão 7).

No tratamento dos dados recolhidos utilizou-se a estatística descritiva (cálculo de frequências) e a estatística inferencial, de forma a tentar detectar as grandes tendências de resposta dos inquiridos, relativamente ao papel que a Internet pode desempenhar na divulgação de informação estatística. Relativamente ao cruzamento de variáveis, o teste de independência do Qui-Quadrado⁵⁸ foi aplicado nos casos em que os pressupostos eram satisfeitos. Como a amostra final é reduzida, os pressupostos do teste do Qui-Quadrado não são satisfeitos na aplicação da maioria dos cruzamentos de variáveis. Para obter as análises foram utilizados os pacotes de *software SPSS, Statística e SPAD (Système Portable d'Analyse des Données, copyright DECISIA, França)*.

⁵⁸ O Teste de Independência do Qui-Quadrado permite verificar a existência de relações de dependência entre duas variáveis qualitativas. Este teste pressupõe que nenhuma célula da tabela tenha frequência esperada inferior a 1 e que não mais de 20% das células tenham frequência esperada inferior a 5 unidades. Em tabelas 2x2, alguns investigadores consideram ser ainda necessário não existir nenhuma célula com frequência esperada inferior a 5 (Pestana & Gageiro, 2003).

VI.3 Caracterização da Amostra

O perfil dos inquiridos é constituído principalmente pelo sector de actividade (por se tratarem de empresas) e se são, ou não, utilizadores do serviço INFOLINE do INE. Os resultados obtidos foram os seguintes:

Quadro 2 – Caracterização da amostra

Variáveis	Frequência Absoluta	Frequência Relativa (%)
Sector de Actividade⁵⁹		
Actividade Financeira	11	18,6
Indústrias Transformadoras	10	16,9
Produção e Distribuição de Electricidade, de Gás e de Água	3	5,1
Construção	3	5,1
Comércio por Grosso e a Retalho	20	33,9
Restauração	2	3,4
Comunicações	10	16,9
Utilizador do Serviço INFOLINE		
Sim	26	44,1
Não	33	55,9
Total	59	100

Pode-se concluir que a amostra é constituída, na sua maioria, por empresas cujos sectores de actividade são o Comércio por Grosso e a Retalho e Actividade Financeira, totalizando 52,5% das empresas inquiridas, sendo 33,9% referente ao Comércio por Grosso e a Retalho. Da totalidade das empresas inquiridas, dez são Indústrias Transformadoras e dez são empresas de Comunicações. De salientar que a maioria das empresas inquiridas não utiliza o serviço INFOLINE do INE.

Para aprofundar a análise foi feito o cruzamento da variável “utilização do serviço” com cada um dos sectores de actividade, cujos resultados são apresentados no quadro seguinte.

⁵⁹ Classificação das Actividades Económicas segundo o INE:
http://www.ine.pt/prodserv/nomenclaturas/cae_rev2.asp (consultado em 30/01/2004).

Quadro 3 – Tabela de Contingência: Sector de Actividade vs a Utilização do Serviço INFOLINE do INE (TS – Total por Sector de Actividade, TU – Total por Utilização do Serviço, T – Total)

Sector de Actividade		Utilização do Serviço INFOLINE		Total
		Sim	Não	
Actividade Financeira	TS	27,3%	72,7%	100%
	TU	11,5%	24,2%	18,6%
	T	3 (5,1%)	8 (13,6%)	11 (18,6%)
Indústrias Transformadoras	TS	70,0%	30,0%	100%
	TU	26,9%	9,1%	16,9%
	T	7 (11,9%)	3 (5,1%)	10 (16,9%)
Produção e Distribuição de Electricidade, de Gás e de Água	TS	66,7%	33,3%	100%
	TU	7,7%	3,0%	5,10%
	T	2 (3,4%)	1 (1,7%)	3 (5,10%)
Construção	TS	66,7%	33,3%	100%
	TU	7,7%	3,0%	5,10%
	T	2 (3,4%)	1 (1,7%)	3 (5,10%)
Comércio por Grosso e a Retalho	TS	35,0%	65,0%	100%
	TU	26,9%	39,4%	33,9%
	T	7 (11,9%)	13 (22,0%)	20 (33,9%)
Restauração	TS	-	100%	100%
	TU	-	6,1%	3,40%
	T	-	2 (3,4%)	2 (3,40%)
Comunicações	TS	50,0%	50,0%	100%
	TU	19,2%	15,2%	16,9%
	T	5 (8,5%)	5 (8,5%)	10 (16,9%)
Total	TS	44,1%	55,9%	100%
	TU	100%	100%	100%
	T	26 (44,1%)	33 (55,9%)	59 (100%)

A análise descritiva das percentagens do quadro anterior permite constatar que 22% da totalidade das empresas inquiridas são do sector de actividade Comércio por Grosso e a Retalho e não utilizam o serviço INFOLINE do INE. Das empresas inquiridas, apenas 11,9% são utilizadoras do referido serviço e pertencem ao mesmo sector de actividade. Aproximadamente 12% são Indústrias Transformadoras e também não utilizam o serviço. São ainda cerca de 14%, em relação ao total de empresas inquiridas, as empresas que se dedicam à actividade financeira e não utilizam o serviço INFOLINE do INE. Aliás, a maioria das empresas deste sector de actividade não utilizam o serviço em análise (~73%). Importa ainda acrescentar que a maioria das empresas utilizadoras do serviço INFOLINE pertencem aos sectores de actividade Indústrias Transformadoras (~27%) e Comércio por Grosso e a Retalho (~27%).

As variáveis sector de actividade e utilização do serviço INFOLINE do INE servirão também para fazer uma análise bivariada dos resultados, permitindo retirar algumas conclusões relativamente a cada um destes grupos.

Parte B – Análise e Interpretação de Resultados

VI.1 Introdução

Como o questionário implementado possui uma pergunta de filtro relativamente à utilização do serviço INFOLINE do INE, a análise e interpretação de resultados vai ser dividida segundo a filtragem referida. Num primeiro momento serão analisados os dados relativos aos utilizadores do serviço, posteriormente serão analisados os dados que referem a questões colocadas a utilizadores e não utilizadores do serviço.

VI.2 Questionário referente aos Utilizadores do Serviço INFOLINE do INE

VI.2.1 Conhecimento do Serviço INFOLINE

A forma como as empresas tomaram conhecimento do serviço INFOLINE do INE constitui um elemento importante para a promoção do próprio serviço. O gráfico seguinte apresenta os resultados obtidos:

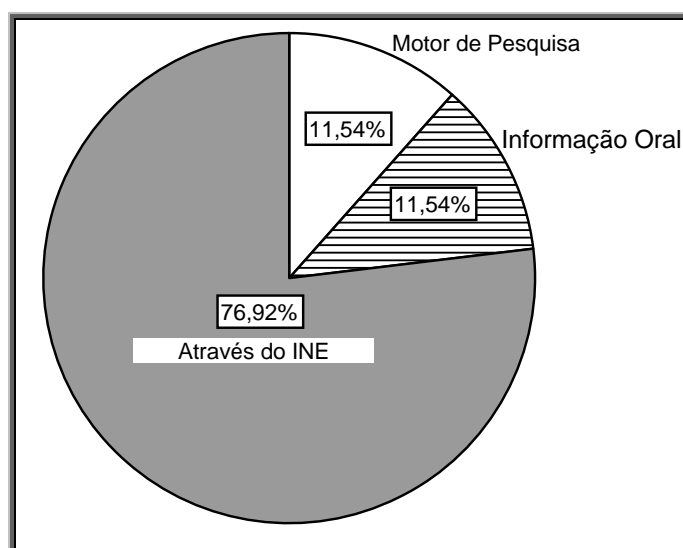


Gráfico 1 – Conhecimento do Serviço INFOLINE

Da observação gráfica, pode-se concluir que a maioria dos utilizadores do serviço INFOLINE (~77%) teve conhecimento deste serviço através do próprio INE. O serviço INFOLINE foi ainda tornado conhecido através da utilização de um motor de busca (12%)

e de comunicação oral (12%). Importa salientar que nenhuma empresa inquirida referiu que conheceu o serviço através da comunicação social, ou de publicações/panfletos, ou de ligações a partir de outras páginas ou, ainda, por *e-mail*.

VI.2.2 Frequência de Utilização do Serviço INFOLINE

Com o aparecimento da *World Wide Web*, a Internet assumiu um papel determinante no acesso a todo o tipo de informação. É a partir deste aparecimento que se dá o crescimento exponencial dos utilizadores da Internet e da quantidade de informação que esta disponibiliza. É a frequência de utilização desta informação, no caso particular do serviço INFOLINE, que se pretende analisar nesta questão.

Quadro 4 – Frequência de Utilização do Serviço INFOLINE

Frequência	Frequência Absoluta	Frequência Relativa (%)
Diariamente	2	7,7
Pelo menos uma por semana	2	7,7
Pelo menos uma vez por mês	17	65,4
Uma vez por trimestre	2	7,7
Uma vez por semestre	3	11,5
Total	26	100

Os primeiros três tipos de frequência (Diariamente, Pelo menos uma vez por semana, Pelo menos uma vez por mês) faziam parte do questionário, contudo o inquirido era convidado a referir um outro tipo de frequência, caso não se identificasse com os tipos apresentados. Os tipos de frequência acrescentados ao questionário foram “Uma vez por trimestre” e “Uma vez por semestre”. A análise do Quadro 4 permite concluir que a maioria das empresas inquiridas e utilizadoras do serviço INFOLINE o utilizam pelo menos uma vez por mês (65,4%). Verifica-se ainda que aproximadamente 12% dos utilizadores do serviço referiram que o utilizam uma vez por semestre e apenas cerca de 8% o utilizam diariamente.

VI.2.3 Tipo de Informação Acedida no Serviço INFOLINE

O serviço INFOLINE do INE disponibiliza vários tipos de informação estatística, conforme analisado no Capítulo IV. Com esta questão, pretende-se averiguar qual o tipo de informação mais (ou menos) acedido. O Quadro 5 apresenta os resultados obtidos.

Quadro 5 – Tipo de Informação acedido no serviço INFOLINE

Tipo de Informação	Acede	Não Acede	Total
Destaques	11 (42,3%)	15 (57,7%)	26 (100%)
Indicadores Estatísticos	24 (92,3%)	2 (7,7%)	
Quadros Estatísticos	11 (42,3%)	15 (57,7%)	
Séries Cronológicas	4 (15,4%)	22 (84,6%)	
Novidades	3 (11,5%)	23 (88,5%)	
Unidades Territoriais	1 (3,8%)	25 (96,2%)	
Informação Rápida	8 (30,8%)	18 (69,2%)	
Retratos Territoriais	1 (3,8%)	25 (96,2%)	
Estudos	9 (34,6%)	17 (65,4%)	

Os resultados que neste campo se obtiveram com o questionário mostram que a maior percentagem é obtida pelos Indicadores Estatísticos (92,3%), seguida pelos Destaques e pelos Quadros Estatísticos, com igual percentagem de 42,3%. Seguem-se ainda os Estudos que são acedidos por aproximadamente 35% dos utilizadores inquiridos e as Folhas de Informação Rápida com 30,8% de utilização. Ao analisarmos a coluna que se refere à percentagem dos tipos de informação que não são acedidos (Quadro 5), verifica-se que a maioria dos utilizadores inquiridos não acedem às Unidades Territoriais e aos Retratos Territoriais, em igual percentagem, 96,2%. De salientar, ainda, que cerca de 89% dos utilizadores não acedem às Novidades e que 84,6% não acedem às Séries Cronológicas.

Com o objectivo de relacionar o sector de actividade com o tipo de informação estatística acedido no serviço INFOLINE, efectuou-se o cruzamento entre estas variáveis (Quadro 6), excluindo o sector de actividade Restauração pelo facto de não existirem utilizadores do serviço INFOLINE deste sector.

Quadro 6 – Tabela de Contingência: Sector de Actividade vs Tipo de Informação Acedido
(A – Acede, NA – Não Acede, T – Total)

Sector de Actividade		Actividade Financeira	Indústrias Transf.	Produção e Distribuição de Electricidade, de Gás e de Água	Construção	Comércio por Grosso e a Retalho	Comunicações	Total
Tipo de Informação	A	2 (7,7%)	4 (15,4%)	-	-	2 (7,7%)	3 (11,5%)	11 (42,3%)
	NA	1 (3,8%)	3 (11,5%)	2 (7,7%)	2 (7,7%)	5 (19,2%)	2 (7,7%)	15 (57,7%)
		T	3 (11,5%)	7 (26,9%)	2 (7,7%)	7 (26,9%)	5 (19,2%)	26 (100%)
Indicadores Estatísticos	A	3 (11,5%)	7 (26,9%)	1 (3,8%)	2 (7,7%)	6 (23,1%)	5 (19,2%)	24 (92,3%)
	NA	-	-	1 (3,8%)	-	1 (3,8%)	-	2 (7,7%)
		T	3 (11,5%)	7 (26,9%)	2 (7,7%)	7 (26,9%)	5 (19,2%)	26 (100%)
Quadros Estatísticos	A	-	3 (11,5%)	2 (7,7%)	-	2 (7,7%)	4 (15,4%)	11 (42,3%)
	NA	3 (11,5%)	4 (15,4%)	-	2 (7,7%)	5 (19,2%)	1 (3,8%)	15 (57,7%)
		T	3 (11,5%)	7 (26,9%)	2 (7,7%)	7 (26,9%)	5 (19,2%)	26 (100%)
Séries Cronológicas	A	-	-	-	-	1 (3,8%)	3 (11,5%)	4 (15,4%)
	NA	3 (11,5%)	7 (26,9%)	2 (7,7%)	2 (7,7%)	6 (23,1%)	2 (7,7%)	22 (84,6%)
		T	3 (11,5%)	7 (26,9%)	2 (7,7%)	7 (26,9%)	5 (19,2%)	26 (100%)
Novidades	A	2 (7,7%)	-	-	-	-	1 (3,8%)	3 (11,5%)
	NA	1 (3,8%)	7 (26,9%)	2 (7,7%)	2 (7,7%)	7 (26,9%)	4 (15,4%)	23 (96,2%)
		T	3 (11,5%)	7 (26,9%)	2 (7,7%)	7 (26,9%)	5 (19,2%)	26 (100%)
Unidades Territoriais	A	-	-	-	-	1 (3,8%)	-	1 (3,8%)
	NA	3 (11,5%)	7 (26,9%)	2 (7,7%)	2 (7,7%)	6 (23,1%)	5 (19,2%)	25 (96,2%)
		T	3 (11,5%)	7 (26,9%)	2 (7,7%)	7 (26,9%)	5 (19,2%)	26 (100%)
Informação Rápida	A	2 (7,7%)	2 (7,7%)	2 (7,7%)	-	-	2 (7,7%)	8 (30,8%)
	NA	1 (3,8%)	5 (19,2%)	-	2 (7,7%)	7 (26,9%)	3 (11,5%)	18 (69,2%)
		T	3 (11,5%)	7 (26,9%)	2 (7,7%)	7 (26,9%)	5 (19,2%)	26 (100%)
Retratos Territoriais	A	-	-	-	-	-	1 (3,8%)	1 (3,8%)
	NA	3 (11,5%)	7 (26,9%)	2 (7,7%)	2 (7,7%)	7 (26,9%)	4 (15,4%)	25 (96,2%)
		T	3 (11,5%)	7 (26,9%)	2 (7,7%)	7 (26,9%)	5 (19,2%)	26 (100%)
Estudos	A	1 (3,8%)	4 (15,4%)	-	1 (3,8%)	2 (7,7%)	1 (3,8%)	9 (34,6%)
	NA	2 (7,7%)	3 (11,5%)	2 (7,7%)	1 (3,8%)	5 (19,2%)	4 (15,4%)	17 (65,4%)
		T	3 (11,5%)	7 (26,9%)	2 (7,7%)	7 (26,9%)	5 (19,2%)	26 (100%)

Da observação do Quadro 6, conclui-se que as Indústrias Transformadoras constituem o sector de actividade que mais consulta os Destaques e os Indicadores Estatísticos.

As empresas que pertencem ao sector Comunicações são as que mais consultam os Quadros Estatísticos, as Séries Cronológicas, e os Retratos Territoriais.

Relativamente às Novidades publicadas no serviço INFOLINE, a maioria das empresas não as consulta, sendo estas, principalmente, dos sectores de actividade Indústrias Transformadoras e Comércio por Grosso e a Retalho.

As empresas que consultam as folhas de Informação Rápida são as que se dedicam à Actividade Financeira, à Produção e Distribuição de Electricidade de Gás e de Água, à área das Comunicações e, por fim, às Indústrias Transformadoras.

No que se refere aos Retratos Territoriais, verifica-se que os sectores de actividade se distribuem de igual forma como no tipo de informação Unidades Territoriais. As empresas utilizadoras do serviço INFOLINE não costumam consultar este tipo de informação, sendo, na sua maioria, dos sectores de actividade Indústrias Transformadoras e Comércio por Grosso e a Retalho.

O tipo de informação publicada nos Estudos cativa os utilizadores do sector de actividade Indústrias Transformadoras apesar de serem apenas 4 das 9 empresas que acedem a esta categoria. Contudo, as empresas dos sectores Comércio por Grosso e a Retalho e Comunicações não utilizam este tipo de informação, sendo, respectivamente, 5 e 4 empresas das 17 que não acedem aos Estudos.

Ao fazer uma análise descritiva do Quadro 6 por sector de actividade conclui-se que as empresas do sector Actividade Financeira costumam consultar os Indicadores Estatísticos, os Destaques, as Novidades e a Informação Rápida. As Indústrias Transformadoras, para além de acederem aos Indicadores Estatísticos, também consultam os Destaques e os Estudos. Estas indústrias revelam ainda pouca importância aos tipos de informação Séries Cronológicas, Novidades, Unidades e Retratos Territoriais.

As empresas que têm como actividade a Produção e Distribuição de Electricidade, de Gás e de Água consultam, normalmente, os Quadros Estatísticos e as folhas de Informação Rápida. Estas empresas apresentam uma fraca utilização dos diversos tipos de informação disponíveis.

Ao nível do sector de actividade Construção, verifica-se que estas empresas acedem aos Indicadores Estatísticos e não revelam muita importância na consulta dos restantes tipos de informação disponíveis no serviço INFOLINE.

O sector de actividade Comércio por Grosso e a Retalho constitui o sector que menos consulta uma maior diversidade de tipos de informação: Destaques, Quadros Estatísticos, Séries Cronológicas, Novidades, Informação Rápida, Retratos Territoriais e

os Estudos; sendo os Indicadores Estatísticos o tipo de informação a que estas empresas acedem. Por último, as empresas de Comunicações consultam praticamente todos os tipos de informação dando mais importância aos Indicadores Estatísticos. Contudo, as empresas deste sector de actividade não acedem aos Retratos Territoriais.

VI.2.4 Temas consultados no Serviço INFOLINE

O conhecimento dos temas que os utilizadores consultam no Serviço INFOLINE constitui um factor importante para perceber qual a área de actuação do utilizador no meio empresarial. O Quadro 7 apresenta as frequências absolutas, e respectivas percentagens, de consultas dos diferentes temas disponibilizados pelo Serviço INFOLINE do INE.

Quadro 7 – Temas consultados no serviço INFOLINE

Temas	Consulta	Não Consulta	Total
Estatísticas Gerais	20 (76,9%)	6 (23,1%)	26 (100%)
Território e Ambiente	3 (11,5%)	23 (88,5%)	
População e Condições Sociais	8 (30,8%)	18 (69,2%)	
Economia e Finanças	24 (92,3%)	2 (7,7%)	
Comércio Externo	8 (30,8%)	18 (69,2%)	
Agricultura, Produção Animal, Silvicultura e Pescas	1 (3,8%)	25 (96,2%)	
Indústria, Construção e Energia	3 (11,5%)	23 (88,5%)	
Comércio Interno, Turismo e Outros Serviços	7 (26,9%)	19 (73,1%)	
Sociedade da Informação e Conhecimento	2 (7,7%)	24 (92,3%)	
Diversos	2 (7,7%)	24 (92,3%)	

A análise descritiva do quadro anterior permite constatar que a maioria das empresas inquiridas e utilizadoras do serviço INFOLINE do INE consultam o tema Economia e Finanças (92,3%). Deve-se acrescentar que aproximadamente 77% dos utilizadores inquiridos consultam o tema Estatísticas Gerais, seguido dos temas População e Condições Sociais e Comércio Externo, em igual percentagem de 30,8%. Os temas menos consultados são a Agricultura, Produção Animal, Silvicultura e Pescas (~96%), Sociedade da Informação e Conhecimento (~92%) e o tema Diversos (~92%).

De igual forma à análise realizada no ponto anterior, irá realizar-se um cruzamento entre as variáveis sector de actividade e temas consultados no serviço INFOLINE do INE (Quadro 8).

Quadro 8 – Tabela de Contingência: Sector de Actividade vs Tema Consultado

(C – Consulta, NC – Não Consulta, T – Total)

Sector de Actividade		Actividade Financeira	Indústrias Transf.	Produção e Distribuição de Electricidade, de Gás e de Água	Construção	Comércio por Grosso e a Retalho	Comunicações	Total
Tipo de Informação	C	2 (7,7%)	5 (19,2%)	2 (7,7%)	-	6 (23,1%)	5 (19,2%)	20 (76,9%)
	NC	1 (3,8%)	2 (7,7%)	-	2 (7,7%)	1 (3,8%)	-	6 (23,1%)
	T	3 (11,5%)	7 (26,9%)	2 (7,7%)	2 (7,7%)	7 (26,9%)	5 (19,2%)	26 (100%)
Território e Ambiente	C	-	1 (3,8%)	1 (3,8%)	-	-	1 (3,8%)	3 (11,5%)
	NC	3 (11,5%)	6 (23,1%)	1 (3,8%)	2 (7,7%)	7 (26,9%)	4 (15,4%)	23 (88,5%)
	T	3 (11,5%)	7 (26,9%)	2 (7,7%)	2 (7,7%)	7 (26,9%)	5 (19,2%)	26 (100%)
População e Condições Sociais	C	-	2 (7,7%)	1 (3,8%)	-	3 (11,5%)	2 (7,7%)	8 (30,8%)
	NC	3 (11,5%)	5 (19,2%)	1 (3,8%)	2 (7,7%)	4 (15,4%)	3 (11,5%)	18 (69,2%)
	T	3 (11,5%)	7 (26,9%)	2 (7,7%)	2 (7,7%)	7 (26,9%)	5 (19,2%)	26 (100%)
Economia e Finanças	C	3 (11,5%)	7 (26,9%)	2 (7,7%)	2 (7,7%)	5 (19,2%)	5 (19,2%)	24 (92,3%)
	NC	-	-	-	-	2 (7,7%)	-	2 (7,7%)
	T	3 (11,5%)	7 (26,9%)	2 (7,7%)	2 (7,7%)	7 (26,9%)	5 (19,2%)	26 (100%)
Comércio Externo	C	1 (3,8%)	2 (7,7%)	-	-	3 (11,5%)	2 (7,7%)	8 (30,8%)
	NC	2 (7,7%)	5 (19,2%)	2 (7,7%)	2 (7,7%)	4 (15,4%)	3 (11,5%)	18 (69,2%)
	T	3 (11,5%)	7 (26,9%)	2 (7,7%)	2 (7,7%)	7 (26,9%)	5 (19,2%)	26 (100%)
Agricultura, Prod. Animal, Silvicultura e Pescas	C	-	1 (3,8%)	-	-	-	-	1 (3,8%)
	NC	3 (11,5%)	6 (23,1%)	2 (7,7%)	2 (7,7%)	7 (26,9%)	5 (19,2%)	25 (96,2%)
	T	3 (11,5%)	7 (26,9%)	2 (7,7%)	2 (7,7%)	7 (26,9%)	5 (19,2%)	26 (100%)
Indústria Construção e Energia	C	-	1 (3,8%)	-	1 (3,8%)	1 (3,8%)	-	3 (11,5%)
	NC	3 (11,5%)	6 (23,1%)	2 (7,7%)	1 (3,8%)	6 (23,1%)	5 (19,2%)	23 (88,5%)
	T	3 (11,5%)	7 (26,9%)	2 (7,7%)	2 (7,7%)	7 (26,9%)	5 (19,2%)	26 (100%)
Comércio Interno, Turismo e Outros Serviços	C	1 (3,8%)	2 (7,7%)	-	-	3 (11,5%)	1 (3,8%)	7 (26,9%)
	NC	2 (7,7%)	5 (19,2%)	2 (7,7%)	2 (7,7%)	4 (15,4%)	4 (15,4%)	19 (73,1%)
	T	3 (11,5%)	7 (26,9%)	2 (7,7%)	2 (7,7%)	7 (26,9%)	5 (19,2%)	26 (100%)
Sociedade da Informação e Conhecimento	C	1 (3,8%)	-	-	-	-	1 (3,8%)	2 (7,7%)
	NC	2 (7,7%)	7 (26,9%)	2 (7,7%)	2 (7,7%)	7 (26,9%)	4 (15,4%)	24 (92,3%)
	T	3 (11,5%)	7 (26,9%)	2 (7,7%)	2 (7,7%)	7 (26,9%)	5 (19,2%)	26 (100%)
Diversos	C	-	1 (3,8%)	-	1 (3,8%)	-	-	2 (7,7%)
	NC	3 (11,5%)	6 (23,1%)	2 (7,7%)	1 (3,8%)	7 (26,9%)	5 (19,2%)	24 (92,3%)
	T	3 (11,5%)	7 (26,9%)	2 (7,7%)	2 (7,7%)	7 (26,9%)	5 (19,2%)	26 (100%)

Da observação do Quadro 8, conclui-se que relativamente aos temas Território e Ambiente, Agricultura, Produção Animal, Silvicultura e Pescas, Indústria, Construção e Energia, Sociedade da Informação e Conhecimento e Diversos, a maioria das empresas não os consulta, sendo estas pertencentes aos sectores de actividade Comércio por Grosso e a Retalho e Indústrias Transformadoras.

Conforme analisado no Quadro 7, os temas Estatísticas Gerais e Economia e Finanças são aqueles que são consultados pela maioria das empresas utilizadoras do serviço INFOLINE que são oriundas de praticamente todos os sectores de actividade. Os sectores de actividade que mais consultam estes temas são Indústrias Transformadoras, Comércio por Grosso e a Retalho e Comunicações.

No que se refere aos temas Território e Ambiente, População e Condições Sociais, Comércio Externo, Comércio Interno, Turismo e Outros Serviços, verifica-se que grande parte das empresas não os consultam sendo estas dos sectores de actividade Indústrias Transformadoras e Comércio por Grosso e a Retalho.

Ao fazer uma análise descritiva do Quadro 8 por sector de actividade conclui-se que as empresas do sector Actividade Financeira costumam consultar os temas Economia e Finanças, que é relativo à sua área e, também, Estatísticas Gerais. Estas empresas revelam ainda pouco interesse pelos temas Território e Ambiente, População e Condições Sociais, Agricultura, Produção Animal, Silvicultura e Pescas e Indústria, Construção e Energia e Diversos.

As Indústrias Transformadoras, para além de acederem ao tema Economia e Finanças, também consultam as Estatísticas Gerais.

As empresas que têm como actividade a Produção e Distribuição de Electricidade, de Gás e de Água consultam, normalmente, as Estatísticas Gerais e Economia e Finanças. Estas empresas apresentam ainda uma fraca consulta aos diversos temas disponíveis no serviço INFOLINE.

Ao nível do sector de actividade Construção, verifica-se que estas empresas acedem ao tema Economia e Finanças e, de igual forma ao sector de actividade anterior, não revelam muita importância na consulta dos restantes temas.

No caso do sector de actividade Comércio por Grosso e a Retalho, este consulta os temas Estatísticas Gerais e Economia e Finanças. Este sector apresenta uma menor consulta numa grande diversidade de temas: Território e Ambiente, Agricultura, Produção Animal, Silvicultura e Pescas, Indústria, Construção e Energia, Sociedade da Informação e Conhecimento e Diversos.

As empresas de Comunicações revelam interesse em consultar os temas Estatísticas Gerais e Economia e Finanças e não consultam os temas Produção Animal, Silvicultura e Pescas, Indústria, Construção e Energia e Diversos.

VI.2.5 Apreciação do Serviço INFOLINE

Nesta questão era pedido aos inquiridos que fizessem uma apreciação, relativamente a alguns aspectos, do serviço INFOLINE. O tipo de resposta é dado numa escala de respostas numéricas, para cada aspecto, em que apenas o significado dos extremos é explícito: 1 (Muito Mau) a 6 (Muito Bom) (Ghiglione & Matalon, 1993).

Quadro 9 – Frequências Relativas e Relativas Acumuladas para a Apreciação do Serviço INFOLINE

Aspectos \ Escala		Muito Mau (1)	(2)	(3)	(4)	(5)	Muito Bom (6)
Organização da Informação (A1)	Freq. Rel.	-	3,8%	11,5%	61,5%	23,1%	-
	Freq. Rel. Ac.	-	3,8%	15,4%	76,9%	100%	-
Quantidade e Diversidade de Informação (A2)	Freq. Rel.	-	-	26,9%	46,2%	26,9%	-
	Freq. Rel. Ac.	-	-	26,9%	73,1%	100%	-
Interactividade (A3)	Freq. Rel.	-	11,5%	34,6%	38,5%	15,4%	-
	Freq. Rel. Ac.	-	11,5%	46,2%	84,6%	100%	-
Design (A4)	Freq. Rel.	-	-	30,8%	53,8%	15,4%	-
	Freq. Rel. Ac.	-	-	30,8%	84,6%	100%	-
Sistema de Navegação (A5)	Freq. Rel.	-	-	30,8%	50,0%	15,4%	3,8%
	Freq. Rel. Ac.	-	-	30,8%	80,8%	96,2%	100%
Conteúdos (A6)	Freq. Rel.	-	-	11,5%	50,0%	38,5%	-
	Freq. Rel. Ac.	-	-	11,5%	61,5%	100%	-
Actualização de Conteúdos (A7)	Freq. Rel.	-	19,2%	7,7%	50,0%	19,2%	3,8%
	Freq. Rel. Ac.	-	19,2%	26,9%	76,9%	96,2%	100%

A análise do quadro anterior permite concluir que a maioria dos utilizadores consideram a *Organização da Informação* no serviço INFOLINE positiva (isto é, igual ou superior a 4) (84,6%), apesar de existir uma percentagem de 15,4% que atribuem uma classificação negativa (isto é, igual ou inferior a 3). Relativamente à *Quantidade e Diversidade de Informação*, 26,9% dos utilizadores do serviço atribuem a classificação de

grau 3, enquanto que os restantes atribuem uma classificação favorável, 46,2% (4) e 26,9% (5).

No que se refere aos aspectos do *Website* INFOLINE, nomeadamente à *Interactividade*, ao *Design* e ao *Sistema de Navegação*, verifica-se que os utilizadores classificam, de uma forma geral, positivamente. Contudo, aproximadamente 46% dos utilizadores atribuem uma classificação negativa à *Interactividade*, 30,8% ao *Design* e 30,8% ao *Sistema de Navegação*.

Ao nível dos *Conteúdos* e respectiva *Actualização*, os inquiridos classificam o primeiro como positivo (88,5%) mas, evidenciam algum descontentamento relativamente à sua actualização. Aproximadamente 27% dos utilizadores referem negativamente este aspecto apesar de 23% o classificarem com os 2 graus mais altos da escala.

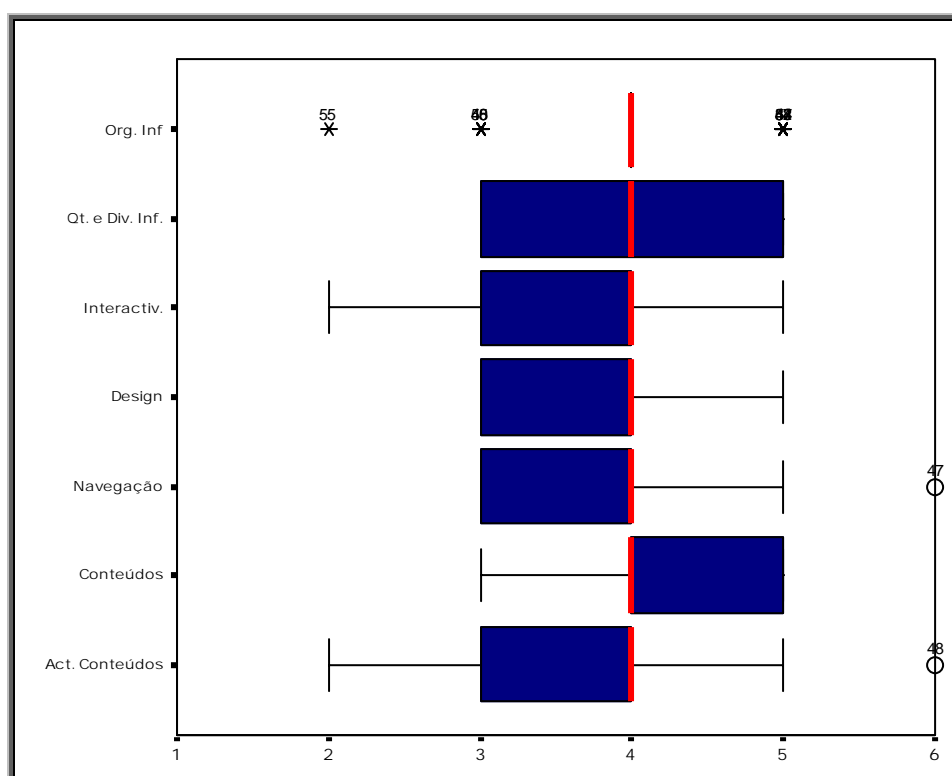


Gráfico 2 – Diagramas de Extremos e Quartis das variáveis que constituem a avaliação do Serviço INFOLINE

No sentido de complementar e reforçar as conclusões anteriores foi construído o Gráfico 2, constituído pelos diagramas de extremos e quartis relativos a cada um dos aspectos avaliados. De facto, a maioria dos inquiridos classifica positivamente o serviço, porque em todos os aspectos o valor da mediana é 4, ou seja, pelo menos 50% dos utilizadores atribuíram uma classificação igual ou superior a 4, a cada uma das características. Verifica-se ainda que nenhum inquirido classificou com *Muito Mau* os

aspectos apresentados e que apenas o *Sistema de Navegação* e a *Actualização dos Conteúdos* tiveram classificação de *Muito Bom* (dada apenas por um indivíduo, em cada um).

A Figura 48 apresenta os diagramas de barras de frequências relativas (%) e permite visualizar, graficamente, a apreciação realizada pelos inquiridos aos diferentes aspectos do serviço INFOLINE. Os diagramas de barras permitem verificar que em todos os aspectos a moda é de 4, ou seja a apreciação é positiva.

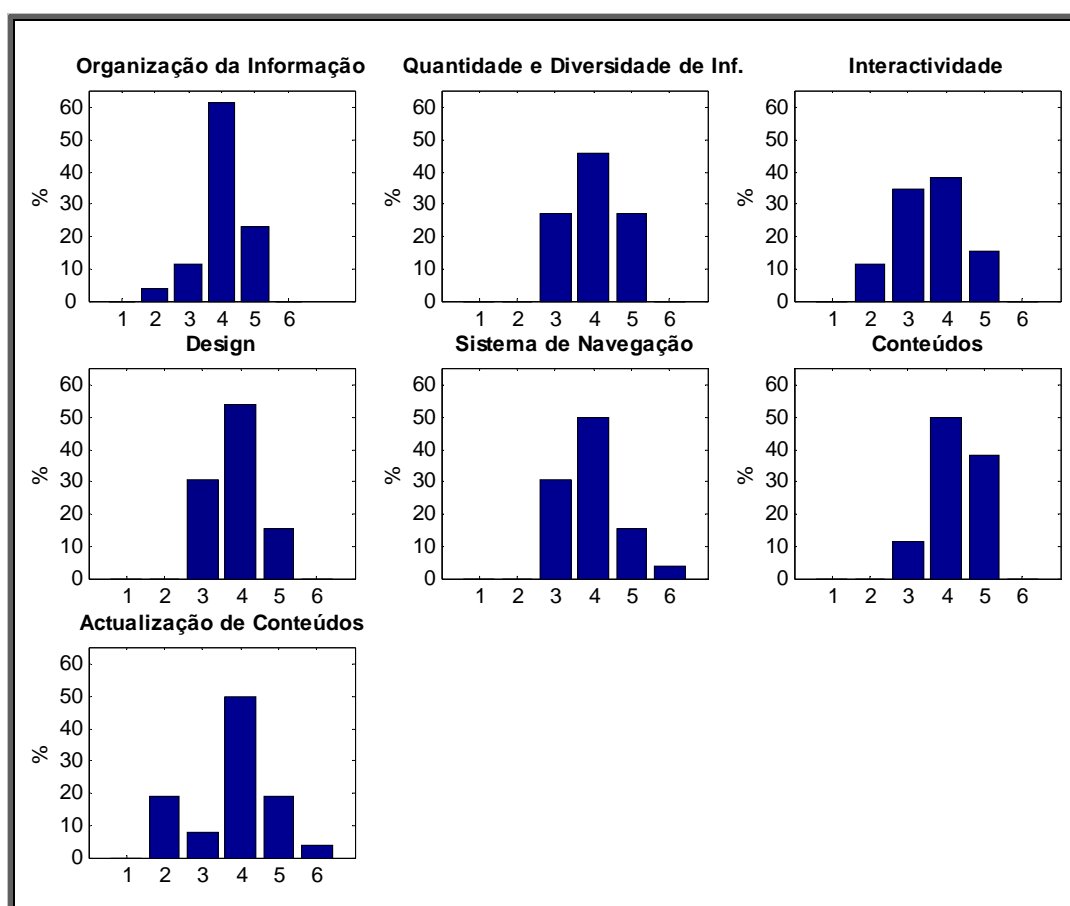


Figura 48 – Diagramas de Barras dos Aspectos de Avaliação

VI.2.5.1 Classificação da Apreciação

À análise realizada no ponto anterior acrescentar-se-á uma Análise Classificatória da apreciação do serviço INFOLINE do INE. Esta análise consistirá na classificação das variáveis relativas à questão 7: A1 – Organização da Informação, A2 – Quantidade e Diversidade de Informação, A3 – Interactividade, A4 – Design, A5 – Sistema de Navegação, A6 – Conteúdos e A7 – Actualização de Conteúdos.

A Análise Classificatória é uma técnica exploratória de análise multivariada que permite agrupar indivíduos ou variáveis em classes homogêneas ou compactas relativamente a uma ou mais características comuns. Cada observação pertencente a uma determinada classe é similar a todas as outras pertencentes a essa classe, e é diferente das observações pertencentes às outras classes (Maroco, 2003).

Como a Análise Classificatória, neste estudo, tem como objectivo o agrupamento de variáveis (e não de indivíduos) e utilizando o *software Statistica*, a medida de semelhança apropriada é o coeficiente de correlação de *Pearson*; observa-se que as sete variáveis de apreciação (A1-A7) são todas positivamente correlacionadas (Quadro 10). O método de agregação hierárquico utilizado neste estudo foi o Método de *Ward*.

Quadro 10 – Coeficiente de Correlação de Pearson entre as variáveis relativas à questão 7

	A1	A2	A3	A4	A5	A6	A7
A1	1,00	0,37	0,46	0,26	0,21	0,23	0,62
A2	0,37	1,00	0,47	0,40	0,60	0,56	0,39
A3	0,46	0,47	1,00	0,55	0,34	0,33	0,36
A4	0,26	0,40	0,55	1,00	0,80	0,45	0,07
A5	0,21	0,60	0,34	0,80	1,00	0,42	0,03
A6	0,23	0,56	0,33	0,45	0,42	1,00	0,35
A7	0,62	0,39	0,36	0,07	0,03	0,35	1,00

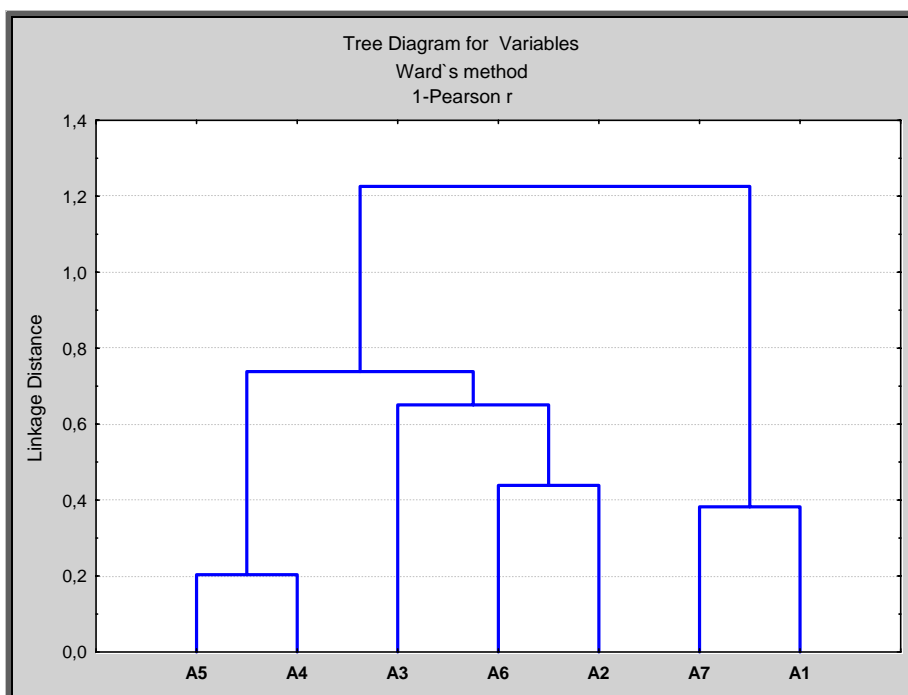


Figura 49 – Dendrograma segundo o Método Ward

Efectuou-se a classificação das sete variáveis de apreciação com base nas respostas dos 26 inquiridos; a hierarquia obtida é representada por um dendrograma. (Figura 49).

A inspecção gráfica do Dendrograma (Figura 49) permite sugerir a existência de 2 classes de variáveis:

- a primeira, constituída pelas variáveis A5, A4, A3, A6 e A2 ou seja, Sistema de Navegação, Design, Interactividade, Conteúdos e Quantidade e Diversidade de Informação, denominada por **Estrutura do Serviço**;
- a segunda, constituída pelas variáveis A7 e A1, ou seja, Actualização de Conteúdos e Organização da Informação, denominada por **Dinamismo do Serviço**.

Contudo, ao efectuar uma inspecção gráfica mais detalhada ao Dendrograma (Figura 49), pode-se sugerir agora uma partição em 4 classes de variáveis:

- a primeira, constituída pelas variáveis A5 e A4, ou seja, Sistema de Navegação e Design denominada por **Interface/Acessibilidade**;
- a segunda, constituída pelas variáveis A7 e A1, ou seja, Actualização de Conteúdos e Organização da Informação, denominada por **Dinamismo do Serviço**;
- a terceira, constituída pelas variáveis A6 e A2, ou seja, Conteúdos e Quantidade e Diversidade de Informação, denominada por **Informação**;
- a quarta constituída apenas pela variável A3, Interactividade, denominada por **Interactividade**.

Relativamente a esta partição, a primeira classe reflecte a opinião dos inquiridos sobre o/a Interface/Acessibilidade que o serviço INFOLINE possui. Analisando a Figura 48 e comparando os diagramas de barras relativos aos aspectos que constituem esta classe, verifica-se que todos os utilizadores os classificaram com uma pontuação igual ou superior a 3, sendo 30,8% a percentagem relativa à pontuação 3.

A segunda classe resume o Dinamismo do serviço INFOLINE com uma pontuação superior ou igual a 2. Contudo, mais de 70% das respostas dos utilizadores é superior ou igual a 4.

A terceira classe representa a Informação do serviço e é classificada, pela maioria dos utilizadores (mais de 70%), com uma pontuação superior ou igual a 4. Neste caso, todos os utilizadores atribuíram uma classificação igual ou superior a 3. Contudo, o

aspecto de avaliação Quantidade e Diversidade de Informação tem uma percentagem superior de classificação 3 (26,9%) do que o aspecto Conteúdos (11,5%).

Por último, a quarta classe destaca-se pela capacidade de intervenção do utilizador durante a utilização do serviço INFOLINE, a Interactividade. A opinião dos utilizadores, relativamente a este aspecto, encontra-se um pouco dividida, já que o classificaram com pontuações entre 2 e 5 inclusive, de uma forma simétrica. Contudo a maioria (cerca de 54%) atribuiu uma pontuação superior ou igual a 4.

VI.2.6 Informação Disponibilizada no Serviço INFOLINE

Segundo Cooke (1999) existem vários critérios de avaliação de uma base de dados, tais como:

- ➔ saber qual a sua cobertura;
- ➔ que áreas temáticas são abordadas;
- ➔ se para cada uma dos temas a informação é completa;
- ➔ que quantidade de informação é fornecida;
- ➔ se a quantidade de informação é suficiente para as necessidades do consumidor.

Com base nestes critérios, pretende-se, com esta questão, averiguar se os utilizadores do serviço INFOLINE encontram toda a informação estatística que procuram. No caso de não encontrarem era pedido que indicassem qual o tipo de informação que gostariam de encontrar.

Quadro 11 – Encontra a informação que procura no serviço INFOLINE

Encontra a Informação	Frequência Absoluta	Frequência Relativa
Sim	24	92,3%
Não	2	7,7%
Total	26	100%

A análise descritiva do Quadro 11 permite concluir que a maioria das empresas utilizadoras do serviço INFOLINE encontra a informação que procura (92,3%). Das empresas inquiridas e utilizadoras do serviço, apenas duas afirmaram que não encontravam a informação pretendida. Os tipos de informação que estes utilizadores referiram que gostariam de encontrar no INFOLINE consistem em informação actualizada e utilizável em base de dados. Contudo, pela análise do serviço realizada no Capítulo IV, verificou-se que o INFOLINE possui uma actualização sistemática de informação. Em

relação ao formato de *download*, o inquirido utilizador não é específico na sua resposta uma vez que o INFOLINE permite a importação de informação estatística em formato de folha de cálculo.

VI.2.7 Análise Multivariada

A análise efectuada foi complementada por uma análise multivariada, consistindo numa análise factorial das correspondências múltiplas seguida de classificação ascendente hierárquica sobre as coordenadas factoriais, retendo-se e interpretando-se a partição mais relevante resultante de “corte” da hierarquia obtida. Para esta análise foram apenas consideradas as variáveis da questão 7 relativa à apreciação do serviço INFOLINE (Organização da Informação, Quantidade e Diversidade de Informação, Interactividade, Design, Sistema de Navegação, Conteúdos, Actualização de Conteúdos) e a variável da questão 4, frequência de utilização.

A análise factorial das correspondências múltiplas identificou um primeiro factor que opõe as empresas que, em relação aos vários aspectos considerados, têm apreciações globalmente mais favoráveis às que exprimem posições mais desfavoráveis. Observa-se, ao longo deste factor, uma gradação de apreciações mais desfavoráveis, num extremo, para apreciações mais favoráveis, no extremo oposto. Os aspectos "a2 – Quantidade e Diversidade de Informação", "a3 – Interactividade" e "a5 – Sistema de Navegação" são os que mais contribuem para a sua formação. O segundo factor parece traduzir uma oposição entre posições mais extremas (pontuações 2 e 5) e posições intermédias (pontuações 3 e 4) (Figura 50).

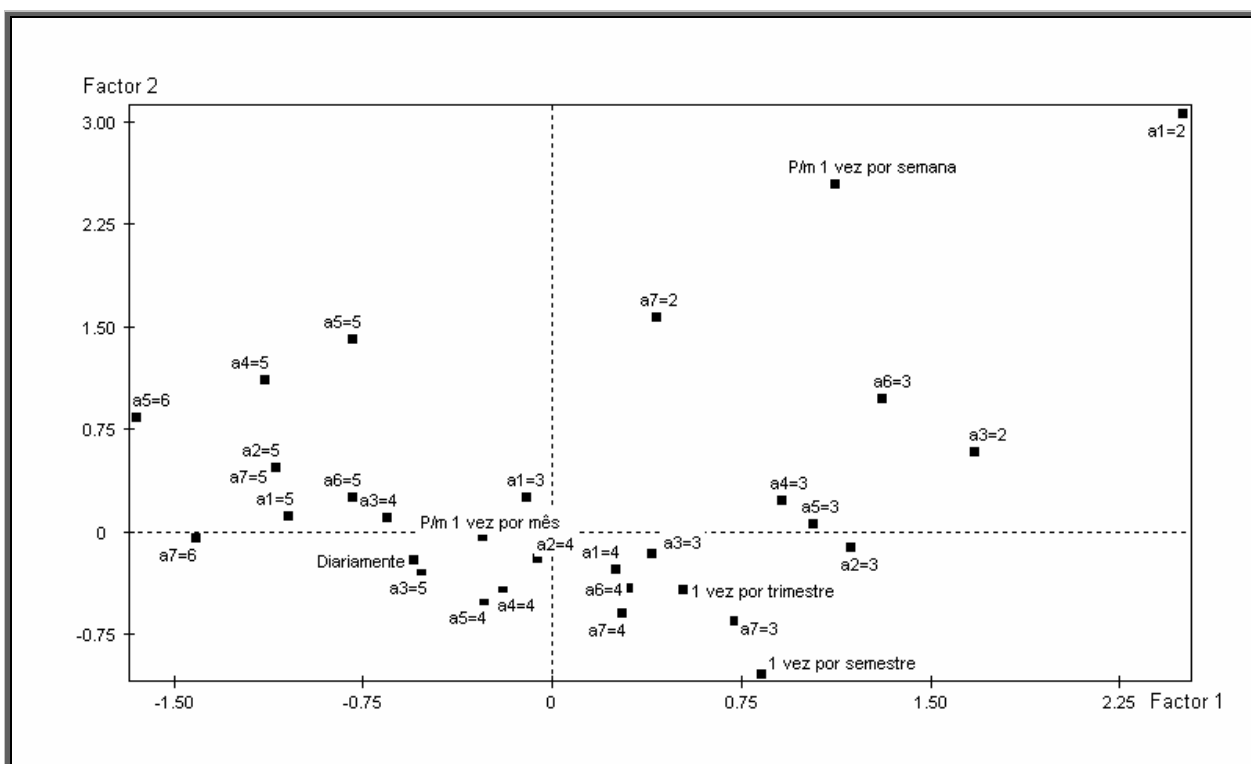


Figura 50 – Primeiro Plano factorial da Análise Factorial de Correspondências Múltiplas das variáveis Apreciação do Serviço e Frequência de Utilização (Ex.: a4=5 significa aspecto 4 com pontuação 5)

A partir das coordenadas dos indivíduos nos primeiros 10 factores – que acumulam cerca de 85% de inércia explicada – efectuou-se então uma classificação ascendente hierárquica dos inquiridos, usando o critério de agregação de Ward. Da árvore obtida (ver Figura 51), foi seleccionada a partição em 3 classes, as quais foram posteriormente consolidadas, de modo a melhorar a sua coesão interna.

Nesta partição que designaremos por $C = \{C1, C2, C3\}$, a classe C1 compreende 16 inquiridos (61,54%), a classe C2, 1 inquirido (3,85%) e a classe C3, 9 inquiridos (34,62%). Da sua descrição em termos das respostas dadas às questões pelos elementos que as compõem, observamos que os elementos da classe C3 tendem a exprimir apreciações mais favoráveis, e os elementos da classe C1 apreciações intermédias. A classe C2 é constituída por um único elemento, que se caracteriza por exprimir opiniões desfavoráveis. Salientemos que este inquirido – número 22 – apresenta uma contribuição de cerca de 24% para a formação do primeiro factor e de cerca de 36% para a formação do segundo factor. A Figura 52 representa a partição projectada no primeiro plano factorial.

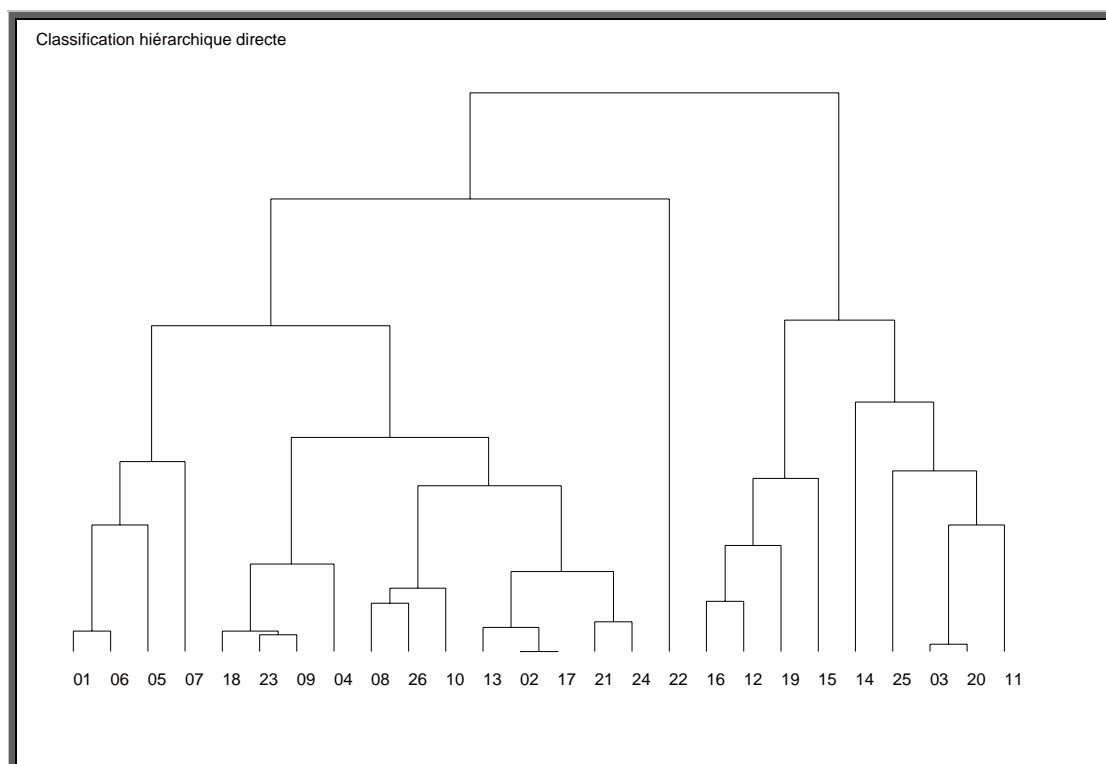


Figura 51 – Dendograma da análise hierárquica ascendente dos inquiridos

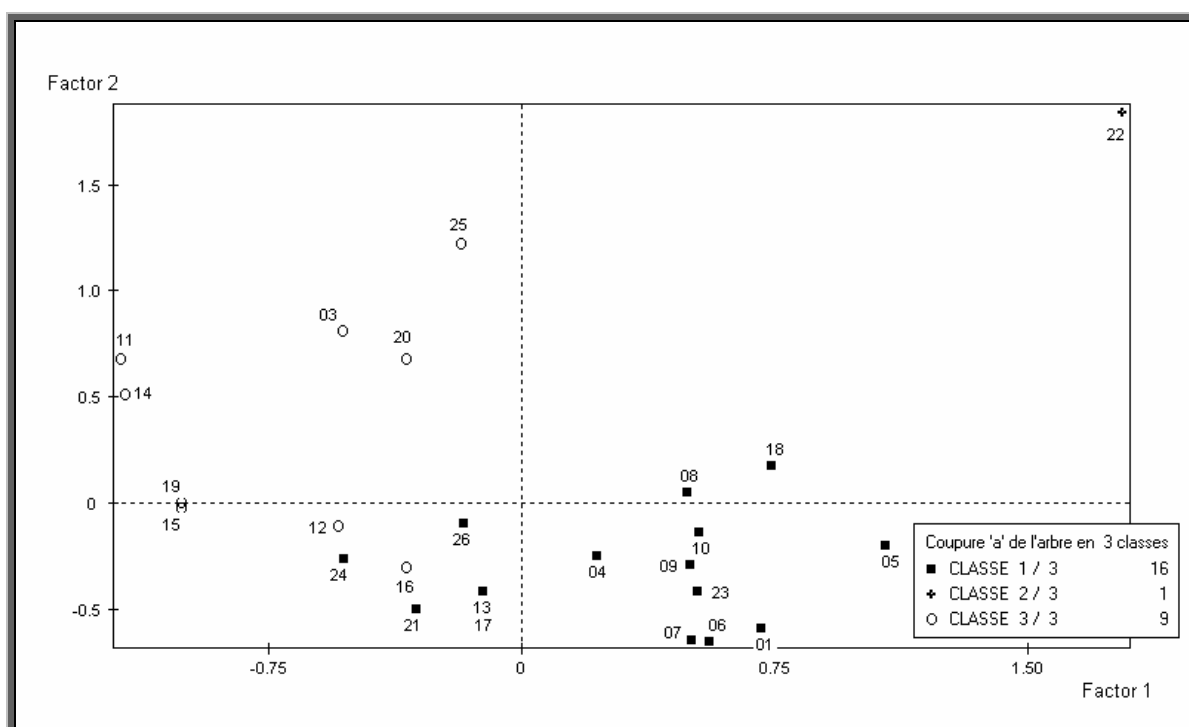


Figura 52 – Representação da partição em três classes no primeiro plano factorial da Análise Factorial de Correspondências Múltiplas dos indivíduos

VI.3 Questionário referente aos Utilizadores e Não Utilizadores do Serviço INFOLINE do INE

O questionário referente a todas as empresas inquiridas, independentemente de serem ou não utilizadoras do serviço INFOLINE do INE, resume-se à questão número nove que pretende averiguar se as empresas recorrem a outros sites de divulgação de informação estatística. Pretende-se ainda saber, no caso de responderem afirmativamente à questão, quais os *sites* a que acedem para consulta de informação estatística e, no caso de não acederem, qual o motivo de não utilizarem este tipo de *sites*.

VI.3.1 Acesso a outros *Sites* de Divulgação de Informação Estatística

O Gráfico 3 representa os resultados obtidos na questão “Acede a outros *sites* de informação estatística/divulgação de dados estatísticos?”.

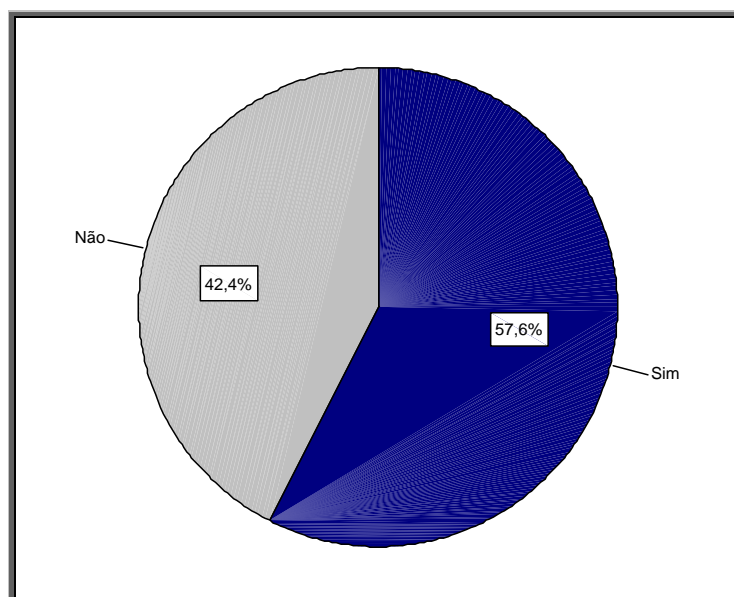


Gráfico 3 – Acesso a outros *Sites* de Divulgação de Informação Estatística

O Gráfico 3 permite concluir que a maioria das empresas acede a outros *sites* de divulgação de informação estatística.

Contudo, seria interessante analisar o cruzamento desta variável com o facto das empresas serem utilizadoras ou não do serviço INFOLINE (Quadro 12).

Quadro 12 – Tabela de Contingência: Utilização do Serviço vs Acesso a outros *Sites* de Informação Estatística

Acesso a outros Sites Utilizador do INFOLINE	Sim	Não	Total
Sim	22 (37,3%)	4 (6,8%)	26 (44,1%)
Não	12 (20,3%)	21 (35,6%)	33 (55,9%)
Total	34 (57,6%)	25 (42,4%)	59 (100%)

A análise descritiva do Quadro 12 permite verificar que a maioria das empresas que utilizam o serviço INFOLINE também recorrem a outros *sites* de divulgação de informação estatística. Esta análise permite concluir ainda, que da totalidade das empresas inquiridas, aproximadamente 36% das empresas não são utilizadoras do serviço INFOLINE e, também não acedem a outros *sites* que divulguem dados estatísticos.

No sentido de complementar a análise anterior pretende-se saber se o acesso a outros *sites* de divulgação de informação estatística é independente do facto de a empresa ser ou não utilizadora do serviço INFOLINE. Para testar esta hipótese, irá aplicar-se o teste de independência do Qui-Quadrado.

Quadro 13 – Teste de Independência do Qui-Quadrado para a relação Utilizador do Serviço INFOLINE/Acesso a outros *Sites* de Informação Estatística (SPSS)

Chi-Square Tests					
	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	13,866 ^b	1	,000		
Continuity Correction ^a	11,960	1	,001		
Fisher's Exact Test				,000	,000
N of Valid Cases	59				

a. Computed only for a 2x2 table

b. 0 cells (,0%) have expected count less than 5. The minimum expected count is 11,02.

Observando o valor da prova 0,000, associado ao teste de independência do Qui-Quadrado (Quadro 13, *Asymp. Sig. (2-sided)*), conclui-se que se rejeita a hipótese de que as variáveis sejam independentes ($p < 0,01$). O Quadro 13 apresenta ainda o resultado do Teste Exacto de Fisher cujo valor de prova é idêntico ao do Qui-Quadrado, corroborando a rejeição da hipótese. Assim, o acesso, ou não, a outros *sites* de divulgação de

informação estatística é influenciado pelo facto de a empresa ser ou não utilizadora do serviço INFOLINE. De facto, como foi analisado anteriormente, os utilizadores do serviço INFOLINE também acedem a outros *sites*; o que poderá ser justificado pela falta de informação que procuram no serviço ou então, por necessitarem de informação mais específica relacionada com o seu sector de actividade, como se irá verificar na análise do ponto seguinte.

Para averiguar quais os sectores de actividade que acedem a outros *sites* de divulgação de informação estatística foi efectuado o cruzamento entre estas duas variáveis (Quadro 14).

Quadro 14 – Tabela de Contingência: Sector de Actividade vs Acesso a outros *Sites* de Informação Estatística (TS – Total por Sector de Actividade, TU – Total por Utilização do Serviço, T – Total)

Sector de Actividade		Acesso a outros <i>Sites</i>		Total
		Sim	Não	
Actividade Financeira	TS	45,5%	54,5%	100%
	TU	14,7%	24,0%	18,6%
	T	5 (8,5%)	6 (10,2%)	11 (18,6%)
Indústrias Transformadoras	TS	60,0%	40,0%	100%
	TU	17,6%	16,0%	16,9%
	T	6 (10,2%)	4 (6,8%)	10 (16,9%)
Produção e Distribuição de Electricidade, de Gás e de Água	TS	66,7%	33,3%	100%
	TU	5,9%	4,0%	5,1%
	T	2 (3,4%)	1 (1,7%)	3 (5,1%)
Construção	TS	66,7%	33,3%	100%
	TU	5,9%	4,0%	5,1%
	T	2 (3,4%)	1 (1,7%)	3 (5,1%)
Comércio por Grosso e a Retalho	TS	60,0%	40,0%	100%
	TU	35,3%	32,0%	33,9%
	T	12 (20,3%)	8 (13,6%)	20 (33,9%)
Restauração	TS	-	100%	100%
	TU	-	8,0%	3,4%
	T	-	2 (3,4%)	2 (3,4%)
Comunicações	TS	70,0%	30,0%	100%
	TU	20,6%	12,0%	16,9%
	T	7 (11,9%)	3 (5,1%)	10 (16,9%)
Total	TS	57,6%	42,4%	100%
	TU	100%	100%	100%
	T	34 (57,6%)	25 (42,4%)	59 (100%)

A análise descritiva do Quadro 14 permite concluir que aproximadamente 20% das empresas inquiridas acedem a outros *sites* de informação estatística e são do sector de actividade Comércio por Grosso e a Retalho. Verifica-se ainda que 11,9% das empresas são da área de Comunicações e acedem a outros *sites* de informação estatística. Importa ainda referir que aproximadamente 14% das empresas não acedem a

outros *sites* de informação estatística e pertencem ao sector de actividade Comércio por Grosso e a Retalho e que 10,2% exercem uma Actividade Financeira e, também não acedem a outros *sites* de divulgação de informação estatística.

Das empresas que acedem a outros *sites* de informação estatística, a maioria pertencem aos sectores de actividade Comércio por Grosso e a Retalho (35,3%) e Comunicações (20,6%). Interessa ainda referir que a maioria das empresas por sector de actividade, à excepção do sector Actividade Financeira, acedem a outros *sites* de informação estatística.

VI.3.1.1 *Sites* de Divulgação de Informação Estatística a que as Empresas Acedem

Nesta questão, as empresas que responderam afirmativamente à pergunta de acesso a outros *sites* de divulgação de informação estatística, teriam de indicar qual ou quais *sites* a que acedem. Como se tratava de uma questão aberta e devido à elevada diversidade de *sites* referidos, foi necessário fazer uma classificação por tipo de *site*, da qual resultaram as seguintes categorias:

- Ministérios;
- Banco de Portugal;
- *Sites* relacionados com o sector de actividade;
- *Sites* de Institutos Estatísticos estrangeiros.

Quadro 15 – Outros *Sites* de Informação Estatística Acedidos

<i>Sites</i> Acedidos	Freq. Absoluta	Freq. Relativa	Total
Ministérios	5	14,7%	34 (100%)
Banco de Portugal	10	29,4%	
<i>Sites</i> relacionados com o sector de actividade	12	35,3%	
<i>Sites</i> de Institutos estrangeiros	7	20,6%	

A análise descritiva do Quadro 15 permite concluir que 35,3% das empresas que afirmaram aceder a outros *sites* de informação estatística, acedem a *sites* relacionados com o seu sector de actividade. Pode-se verificar ainda que aproximadamente 30% das empresas que utilizam outros *sites* de divulgação de informação estatística acedem ao Banco de Portugal e que 7 empresas acedem a outros Institutos de Estatística.

Perante esta perspectiva seria interessante relacionar o sector de actividade com o tipo de *sítes* acedidos. O quadro seguinte resume os resultados obtidos:

Quadro 16 – Tabela de Contingência: Sector de Actividade vs Outros *Sítes* acedidos (TS – Total por Sector de Actividade, TU – Total por Utilização do Serviço, T – Total)

Sector de Actividade		Outros <i>sítes</i> acedidos				Total
		Ministérios	<i>Sítes</i> relacionados com o sector de actividade	Banco de Portugal	<i>Sítes</i> de Institutos Estrangeiros	
Actividade Financeira	TS	-	20,0%	60,0%	20,0%	100%
	TU	-	10,0%	25,0%	14,3%	14,7%
	T	-	1 (2,9%)	3 (8,8%)	1 (2,9%)	5 (14,7%)
Indústrias Transformadoras	TS	33,3%	-	50,0%	16,7%	100%
	TU	40,0%	-	25,0%	14,3%	17,6%
	T	2 (5,9%)	-	3 (8,8%)	1 (2,9%)	6 (17,6%)
Produção e Distribuição de Electricidade, de Gás e de Água	TS	-	-	50,0%	50,0%	100%
	TU	-	-	8,3%	14,3%	5,9%
	T	-	-	1 (2,9%)	1 (2,9%)	2 (5,9%)
Construção	TS	-	-	100%	-	100%
	TU	-	-	16,7%	-	5,9%
	T	-	-	2 (5,9%)	-	2 (5,9%)
Comércio por Grosso e a Retalho	TS	25,0%	50,0%	8,3%	16,7%	100%
	TU	60,0%	60,0%	8,3%	28,6%	35,3%
	T	3 (8,8%)	6 (17,6%)	1 (2,9%)	2 (5,9%)	12 (35,3%)
Restauração	TS	-	-	-	-	-
	TU	-	-	-	-	-
	T	-	-	-	-	-
Comunicações	TS	-	42,9%	28,6%	28,6%	100%
	TU	-	30,0%	16,7%	28,6%	20,6%
	T	-	3 (8,8%)	2 (5,9%)	2 (5,9%)	7 (20,6%)
Total	TS	14,7%	29,4%	35,3%	20,6%	100%
	TU	100%	100%	100%	100%	100%
	T	5 (14,7%)	10 (29,4%)	12	7 (20,6%)	34 (100%)

A análise descritiva do Quadro 16 permite concluir que 17,6% das empresas são da área de Comércio por Grosso e a Retalho e acedem a *sítes* relacionados com o seu sector de actividade. Verifica-se ainda que as empresas que consultam o Banco de Portugal pertencem a uma grande variedade de sectores de actividade, exceptuando as empresas do sector Restauração que, pela análise do quadro, não acedem a outros *sítes* de informação estatística. Importa ainda referir que a maioria das empresas do sector Actividade Financeira consultam o Banco de Portugal, assim como as Indústrias Transformadoras. Os *sítes* de institutos estrangeiros são consultados na maioria por empresas dos sectores de actividade Comércio por Grosso e a Retalho (28,6%) e Comunicações (28,6%).

VI.3.1.2 Motivo de Não Utilização de *Sites* de Divulgação de Informação Estatística

Conforme o Gráfico 3, cerca de 42% das empresas que responderam ao questionário não acedem a outros *sites* de divulgação de informação estatística. Nesta questão, os inquiridos eram convidados a indicar qual o motivo de não acederem a este tipo de *sites*. Como se tratava de uma questão aberta, foi necessário fazer uma classificação por motivo, da qual resultaram as seguintes categorias:

- ➔ Irrelevante para o desempenho da actividade;
- ➔ Produção de informação a nível interno;
- ➔ Desconhecimento/credibilidade;
- ➔ Recebimento de publicações especializadas.

Quadro 17 – Motivos de Não Utilização de *sites* de Informação Estatística

Motivo	Freq. Absoluta	Freq. Relativa	Total
Irrelevante para o desempenho da actividade	12	48,0%	25 (100%)
Produção de informação a nível interno	2	8,0%	
Desconhecimento/Credibilidade	3	12,0%	
Recebimento de publicações especializadas	8	32,0%	

O Quadro 17 apresenta os resultados obtidos nesta questão. Verifica-se que 48% das empresas que não acedem a outros *sites* de informação estatística não o fazem por considerar irrelevante para o desempenho das suas actividades. A análise descritiva do Quadro 17 permite concluir ainda que 32% das empresas recebem publicações especializadas e, por isso, não acedem a outros *sites* de informação estatística.

No sentido de averiguar quais os motivos de não acesso a outros *sites* de divulgação de informação estatística que as empresas utilizadoras do serviço INFOLINE do INE apresentam, efectuou-se o cruzamento dessas variáveis (Quadro 18).

Quadro 18 – Tabela de Contingência: Utilizador INFOLINE vs Motivos de Não Utilização de *sítes* de Informação Estatística (TM – Total por Motivo, TU – Total por Utilização do Serviço, T – Total)

Motivo		Utilizador INFOLINE		Total
		Sim	Não	
Irrelevante para o desempenho da actividade	TM	8,3%	91,7%	100%
	TU	25,0%	52,4%	48,0%
	T	1 (4,0%)	11 (44,0%)	12 (48,0%)
Produção de informação a nível interno	TM	-	100,0%	100%
	TU	-	9,5%	8,0%
	T	-	2 (8,0%)	2 (8,0%)
Desconhecimento/Credibilidade	TM	66,7%	33,3%	100%
	TU	50,0%	4,8%	12,0%
	T	2 (8,0%)	1 (4,0%)	3 (12,0%)
Recebimento de publicações especializadas	TM	12,5%	87,5%	100%
	TU	25,0%	33,3%	32,0%
	T	1 (4,0%)	7 (28,0%)	8 (32,0%)
Total	TM	16,0%	84,0%	100%
	TU	100,0%	100,0%	100%
	T	4 (16,0%)	21 (84,0%)	25 (100%)

A análise descritiva do quadro anterior permite concluir que apenas 4 empresas utilizadoras do serviço INFOLINE não acedem a outros *sítes* de divulgação de informação estatística (conforme analisado também no Quadro 12). Destas, duas não utilizam outros *sítes* por desconhecimento ou por não os considerarem credíveis. Do conjunto das empresas não utilizadoras do serviço INFOLINE (21), onze também não utilizam outros *sítes* de divulgação de informação estatística por considerarem que este tipo de informação é irrelevante para as suas actividades. Verifica-se ainda que 28% das empresas são não utilizadoras do serviço INFOLINE e não acedem a outros *sítes* de divulgação de informação estatística porque recebem publicações específicas da área de interesse.

No sentido de complementar a análise anterior, efectuou-se ainda o cruzamento do motivo de não utilização de outros *sítes* de divulgação de informação estatística com o sector de actividade das empresas (Quadro 19).

Quadro 19 – Tabela de Contingência: Sector de Actividade vs Motivos de Não Utilização de *sites* de Informação Estatística

Motivo Sector de Actividade	Irrelevante para o desempenho da actividade	Produção de informação a nível interno	Desconhecimento /Credibilidade	Recebimento de publicações especializadas	Total
Actividade Financeira	1 (4,0%)	2 (8,0%)	-	3 (12,0%)	6 (24,0%)
Indústrias Transformadoras	-	-	2 (8,0%)	2 (8,0%)	4 (16,0%)
Produção e Distribuição de Electricidade, de Gás e de Água	-	-	-	1 (4,0%)	1 (4,0%)
Construção	1 (4,0%)	-	-	-	1 (4,0%)
Comércio por Grosso e a Retalho	6 (24,0%)	-	1 (4,0%)	1 (4,0%)	8 (32,0%)
Restauração	2 (8,0%)	-	-	-	2 (8,0%)
Comunicações	2 (8,0%)	-	-	1 (4,0%)	3 (12,0%)
Total	12 (48,0%)	2 (8,0%)	3 (12,0%)	8 (32,0%)	25 (100%)

Os resultados apresentados no Quadro 19 permitem concluir que 24% das empresas consideradas se dedicam ao Comércio por Grosso e a Retalho e não acedem a outros *sites* de divulgação de informação estatística por considerarem irrelevante para a sua actividade. Verifica-se ainda que 12% das empresas inquiridas são do sector financeiro e não acedem a outros *sites* de informação estatística porque recebem publicações relativas à sua área de actuação no mercado.

VI.4. Conclusões

Como referido anteriormente o questionário implementado foi enviado a duzentas grandes empresas, das quais cinquenta e nove constituíram a amostra final. A amostra caracteriza-se essencialmente por empresas do sector de actividade Comércio por Grosso e a Retalho (33,9%) e por empresas que exercem uma actividade financeira (18,6%). A amostra final caracteriza-se ainda por utilizadores (44,1%) e por não utilizadores (55,9%) do serviço INFOLINE do INE.

Do cruzamento do sector de actividade e utilização do serviço INFOLINE do INE conclui-se que as empresas utilizadoras são na sua maioria Indústrias Transformadoras (~27%) e pertencem ao sector de actividade Comércio por Grosso e a Retalho (~27%).

Por outro lado, a maioria das empresas não utilizadoras do serviço exercem uma Actividade Financeira (~24%) e Comércio por Grosso e a Retalho (~39%).

A análise e interpretação dos resultados foram divididas em duas vertentes. Por um lado, pretendia-se perceber a utilização do serviço e por outro, entender o porquê da não utilização do serviço INFOLINE do INE e/ou das empresas não recorrerem a *sítes* de divulgação de dados estatísticos.

Ao nível dos utilizadores obtiveram-se as seguintes conclusões:

- A maioria dos utilizadores do serviço INFOLINE do INE teve conhecimento deste através do próprio INE. É importante referir que as opções de promoção do serviço Comunicação Social, Publicações/Planfetos, *Link* e *e-mail* não constituirão um meio adequado de divulgação do serviço, já que nenhuma empresa tomou conhecimento do referido serviço através destas.
- A frequência de utilização do serviço por parte das empresas utilizadoras e inquiridas pode-se considerar baixa, uma vez que apenas 2 empresas consultam diariamente o serviço. Contudo, 65,4% dos utilizadores consultam pelo menos uma vez por mês.
- O serviço INFOLINE do INE possui Tipos de Informação aos quais as empresas utilizadores poderão aceder. Assim, verificou-se que 92,3% dos utilizadores acedem aos Indicadores Estatísticos, 42,3% acedem aos Destaques e 42,3% acedem aos Quadros Estatísticos. De referir ainda que os Tipos de Informação menos acedidos pelas empresas utilizadoras são as Unidades Territoriais e os Retratos Territoriais.
- Do cruzamento das variáveis Sector de Actividade e Tipo de Informação a que as empresas acedem, conclui-se que, das empresas que consultam os Indicadores Estatísticos, a maioria são Indústrias Transformadoras e pertencem ao sector de actividade Comércio por Grosso e a Retalho. Das que acedem aos Destaques (11), 4 são Indústrias Transformadoras e 3 pertencem ao sector de actividade Comunicações. Relativamente aos Quadros Estatísticos, verifica-se que das empresas que consultam este Tipo de Informação (11), 4 são do sector de actividade Comunicações.
- A pesquisa de informação estatística no serviço INFOLINE encontra-se dividida por temas estatísticos. Da análise desta questão conclui-se que 92,3% empresas utilizadoras consultam o tema Economia e Finanças e

76,9% Estatísticas Gerais. Os temas estatísticos menos consultados são Agricultura, Produção Animal, Silvicultura e Pescas (96,2%), Sociedade da Informação e Conhecimento (92,3%) e Diversos (92,3%). Embora os temas Território e Ambiente (88,5%) e Indústria, Construção e Outros Serviços (88,5%) tenham igualmente percentagens elevadas de não acesso.

- ➔ Do cruzamento das variáveis Sector de Actividade e Tema Estatístico resulta que das empresas que consultam o tema Economia e Finanças (24), 7 são Indústrias Transformadoras, 5 são do sector de actividade Comércio por Grosso e a Retalho e 5 são da área de Comunicações. Relativamente às Estatísticas Gerais, das 20 empresas que as consultam, 6 pertencem ao sector de actividade Comércio por Grosso e a Retalho.
- ➔ No que se refere à apreciação do serviço INFOLINE do INE, concluiu-se que as empresas utilizadoras o consideram razoável, uma vez que a maioria das empresas utilizadoras atribuiu uma classificação superior ou igual a 4, numa escala com valor mínimo de 1 – Muito Mau – a um valor máximo de 6 – Muito Bom -, a todos os aspectos de avaliação do *site* e serviço. De salientar que o aspecto ao qual atribuíram a maior percentagem de pontuação superior ou igual a 4 foi o Conteúdo (88,5%), seguido do aspecto Organização da Informação (84,6%). Contudo o aspecto que menor percentagem teve, no que diz respeito a uma pontuação igual ou superior a 4, foi a Interactividade (53,9%).
- ➔ A nível da informação disponibilizada no serviço INFOLINE do INE, a maioria das empresas utilizadoras do serviço encontra a informação de que necessita (92,3%).

Ao nível dos utilizadores e não utilizadores retiraram-se as seguintes conclusões:

- ➔ A maioria das empresas acede a outros *sites* de divulgação de informação estatística (57,6%).
- ➔ A maioria das empresas que utilizam o serviço INFOLINE, também recorre a outros *sites* de divulgação de informação estatística. Verificou-se ainda que o acesso, ou não, a outros *sites* de divulgação de informação estatística é dependente do facto de a empresa ser ou não utilizadora do serviço INFOLINE.

- Das empresas que acedem a outros *sites* (34 empresas), 12 pertencem ao sector de actividade Comércio por Grosso e a Retalho e 7 ao de Comunicações. Das empresas que não acedem a outros *sites* de divulgação de informação estatística (25 empresas), 8 são do sector de actividade Comércio por Grosso e a Retalho e 6 exercem actividade financeira.
- A maioria das empresas que acedem a outros *sites* de divulgação de informação estatística consulta *sites* relacionados com o seu sector de actividade (35,3%) e o *site* do Banco de Portugal (29,4%). Foram ainda mencionados como outros *sites* de divulgação de informação estatística, *sites* de Ministérios e *sites* de Institutos Estatísticos estrangeiros.
- Das empresas que acedem ao *site* do Banco de Portugal (12), 3 exercem uma actividade financeira e 3 são Indústrias Transformadoras. A maioria das empresas que acedem a *sites* relacionados com o sector de actividade são do sector de actividade Comércio por Grosso e a Retalho.
- As empresas que não acedem a outros *sites* de divulgação de informação estatística justificam-se por considerarem irrelevante para o desempenho da actividade que exercem (48%) e por receberem de publicações especializadas (32,0%).
- Apenas 4 empresas utilizadoras do serviço INFOLINE do INE não acedem a outros *sites* de divulgação de informação estatística. De referir que 2 destas empresas não utilizam outros *sites*, por desconhecimento ou por não os considerarem credíveis. Verifica-se ainda que 11 empresas não utilizam o serviço INFOLINE e não acedem a outros *sites* de informação estatística por considerarem irrelevante para o exercício das suas actividades.

Por último, a análise factorial das correspondências múltiplas, no caso das variáveis de apreciação do serviço INFOLINE, permitiu perceber a existência de dois factores. Um primeiro factor que opõe as empresas que têm opiniões globalmente mais favoráveis às que exprimem posições mais desfavoráveis, e um segundo factor traduz uma oposição entre posições mais extremas e posições intermédias.

Conclusões, Limitações e Perspectivas

As reflexões operadas ao longo deste trabalho vão de encontro à célebre frase de H. G. Wells⁶⁰: “No futuro, o pensamento estatístico será tão necessário para a cidadania eficiente como saber ler e escrever.”. De facto, as respostas ao inquérito realizado vêm confirmar a ideia base deste trabalho. A análise dos inquéritos permitiu concluir que aproximadamente 64% das empresas inquiridas consultam *sites* de informação estatística (conforme Quadro 12, Capítulo VI) para poderem tomar decisões na gestão das suas empresas, ou seja, os empresários procuram estar actualizados estatisticamente.

A actividade de pesquisa de informação estudada no Capítulo II (Choo, 2000) descreve a forma com que os empresários do mundo actual utilizam a informação como matéria-prima para criar ou aumentar a competitividade das suas empresas. Como se viu, segundo Choo (2000) a actividade de pesquisa de informação divide-se em três momentos. De facto, e extrapolando esta actividade para o caso de informação estatística, tem-se o primeiro momento, definido como necessidade de informação, em que o empresário perante percepções, situações novas, motivos, etc., sente necessidade de fundamentar as suas decisões através de informação estatística. No segundo momento, denominado por Choo (2000) como pesquisa de informação, é onde o empresário tira partido das potencialidades que a Internet oferece ao nível de acesso a informação estatística. Por último, o terceiro momento é aquele em que o empresário utiliza a informação estatística, que obteve no momento anterior, na construção da sua estratégia empresarial.

O facto de o empresário poder aceder a informação estatística na Internet e não apenas em formato papel, revela a importância que esta tem no acesso rápido a este tipo de informação. Os Institutos Nacionais de Estatística assumem aqui um papel preponderante, nomeadamente na produção e, principalmente nas formas de divulgação e apresentação de informação estatística. É importante, no sentido de que oferece uma informação fiável, válida e oportuna e, principalmente, actualizada. Esta informação é sempre acompanhada de sumários que identificam e caracterizam o fenómeno estudado, para além de respeitar o segredo estatístico referido no Capítulo IV.

No que se refere aos serviços de divulgação de informação estatística de institutos estrangeiros e comparando com o INFOLINE, conclui-se que o Instituto

⁶⁰ H. G. Wells, (1866-1946), Romancista inglês, jornalista, sociólogo e historiador, famoso pelo seu trabalho na área de ficção científica. <http://www.online-literature.com/wellshg/> (consultado em 15/04/2005)

Nacional de Estatística português deve preocupar-se em continuar a fornecer sempre uma grande diversidade de temas mas, principalmente, deve apostar na criação de formas inovadoras de divulgação de informação estatística. Viu-se que no serviço STATLINE do Instituto holandês o utilizador poderia “fabricar” *online* a informação estatística pretendida, o que o INFOLINE não oferece.

No que respeita a hipótese de trabalho considerada no início da investigação, parece que se encontra plenamente aceite. Com efeito, nesta hipótese, considerou-se que «*A Internet constitui um acesso inovador a informação estatística e de divulgação da mesma, de modo a dar resposta às necessidades da população empresarial no que se refere à tomada de decisões a esse nível*». Esta hipótese encontra o seu expoente máximo no referido serviço STATLINE do *Statistics Netherlands* que se analisou no Capítulo IV, pela utilização de um programa chamado “*webselector*”. Esta função permite ao utilizador gerar a informação estatística pretendida através da selecção e posterior cruzamento de variáveis. Como este serviço se encontra disponível *online* e de forma gratuita, está ao alcance de qualquer empresário que ao sentir necessidade de uma informação estatística específica, a pode obter de uma forma inovadora, rápida e a qualquer hora do dia.

Um outro tema que esta investigação pretendia analisar era a identificação do grau de satisfação dos utilizadores de um serviço de divulgação de informação estatística, nomeadamente, o INFOLINE do INE. A amostra recolhida caracteriza-se fundamentalmente por empresas do sector de actividade Comércio por Grosso e a Retalho (33,9%) ou por exercerem uma Actividade Financeira (18,6%). Contudo, apenas 26 das empresas inquiridas (59) utilizam o serviço INFOLINE, constituindo assim a amostra de utilizadores. O inquérito permitiu concluir que os utilizadores inquiridos conheceram este serviço através do INE (77%) e utilizam-no pelo menos uma vez por mês (65,4%). As empresas utilizadoras do serviço acedem normalmente a Indicadores Estatísticos (92,3%), Destaques (42,3%) e Quadros Estatísticos (42,3%). A apreciação do serviço INFOLINE, pelas empresas inquiridas, é razoável, uma vez que a maioria dos inquiridos classificaram os aspectos, que descrevem o *site* e o serviço, com um grau igual ou superior a 4 (conforme Gráfico 2, Capítulo VI). Importa ainda referir que a maioria das empresas utilizadoras encontram a informação estatística que procuram, o que revela a grande diversidade de temas que o serviço possui. Como se referiu, a maioria das empresas inquiridas acedem a outros *sites* de informação estatística que não o serviço INFOLINE do INE. O Banco de Portugal foi um dos mais citados (29,4%) como uma boa referência ao nível da informação estatística. Salienta-se ainda que algumas empresas

ainda desconhecem o valor da informação estatística, considerando-a irrelevante para o exercício das suas actividades (12 empresas). Destas, seis empresas pertencem ao sector de Comércio por Grosso e a Retalho e quatro, por utilizarem o serviço INFOLINE, consideram irrelevante aceder a outros *sites*.

Por último, refere-se uma nova abordagem na apresentação e divulgação de informação estatística: a análise de dados simbólicos. Esta nova filosofia de análise multivariada de dados consiste em estender as tradicionais técnicas de análise estatística às variáveis simbólicas, que assumem valores como subconjuntos de categorias, intervalos ou distribuições de frequências, ao contrário da análise clássica onde as variáveis apenas podem assumir um único número ou categoria. Esta metodologia inovadora está em contínuo desenvolvimento e aperfeiçoamento e os Institutos Nacionais devem explorar esta vertente de forma a inovarem na apresentação e divulgação de informação estatística, já que apresenta vantagens claras para as estatísticas oficiais, conforme referido no Capítulo V.

Em síntese, pode-se afirmar que a Internet é um factor de inovação no acesso a informação estatística e que as empresas aderem a este meio de comunicação de forma a ultrapassar o analfabetismo estatístico.

Contudo, o método adoptado na pesquisa deste estudo apresentou algumas limitações nomeadamente no que se refere à recolha de dados. O meio utilizado para a obtenção de dados foi o envio de um questionário por correio, o que traduziu num volume pequeno de respostas. Neste aspecto, seria desejável que fossem empreendidas novas pesquisas com maior vulto e poder de divulgação. Como sugestão, o questionário poderia ser enviado por *e-mail* às empresas, ou poderia ser implementado *online*, disponível ao se aceder à página do Instituto. Este tipo de solução permitiria para além de uma redução de custos e/ou aumento do número e diversidade de inquiridos, poupar, também, tempo a quem responde, que não precisaria de devolver por correio/fax o questionário.

A divulgação de informação estatística em Portugal necessita de um impulso. Impulso esse que leva a um caminho que depende de uma palavra-chave: inovação. Alguns *sites* de divulgação de informação estatística apresentam-se como meros repositórios de dados estatísticos. O instituto que, hoje em dia, recolhe, classifica, sistematiza e analisa, também actualiza o passado a cada momento ao divulgar informação estatística. Assim, terá de se abrir para a sociedade a partir do reconhecimento do seu novo papel de não apenas produtor, mas também de divulgador de dados para pesquisas científicas, educacionais e empresariais.

Porém, não se pode perder de vista a realidade das novas tecnologias da comunicação. Trata-se de saber como colocar os meios de comunicação ao serviço da pesquisa, do público em geral e do público empresarial, sem perder a qualidade e confidencialidade do conteúdo a ser apresentado.

Posicionar os institutos nas tecnologias da informação é construir um instituto dialógico e interactivo, ou seja, deixar cada vez mais em aberto a possibilidade de ser o utilizador a criar a informação desejada, em relação ao que está exposto. Assim, cria-se um espaço para o público decidir, motivando os visitantes, pesquisadores e empresários. Aliás, a interactividade é o factor que mais novidade traz, é uma forma de se aproximar cada vez mais do cidadão e tornar o espaço na Internet mais atraente, mais dinâmico, com uma linguagem e recursos técnicos mais atraentes. Nesse sentido, é preciso repensar os conceitos de transmissão das novas técnicas de pesquisa de informação estatística, através da Internet, colocando-as ao serviço de um pensamento crítico.

No futuro, quem divulga informação estatística através da Internet, já que este meio representa uma revolução comunicativa, terá de o fazer sempre acompanhado das palavras-chave inovação e interactividade. É importante inovar mas principalmente interagir com os utilizadores de forma a motivá-los a utilização, no seu dia-a-dia empresarial, um *site* de divulgação de informação estatística.

Bibliografia

AFUAH, Allan – **Innovation Management. Strategies, Implementations and Profits.** Oxford University Press, 1998.

AGUILLAR, Francis J. – **Scanning the Business Environment.** New York, Macmillan Co., 1967

ALBARELLO, Luc [tal.] – **Práticas e Métodos de Investigação em Ciências Sociais.** Lisboa: Publicações Gradiva, 1997.

ALONSO, José António F. – **A Produção de Informação e Conhecimento.** São Paulo em Perspectiva. Brasil: Fundação de Economia e Estatística, 16 (3), 2002. p. 23-30.

ALVES, Caetano; BANDEIRA, Sílvia – **Dicionário de Marketing.** 2.^a ed. Porto: Edições IPAM, 2001.

ALVES, José Augusto; CAMPOS, Pedro; BRITO, Pedro Quelhas – **O Futuro da Internet, Estado da arte e Tendências de evolução da Internet.** Portugal: Edições Centro Atlântico, 1999.

ARAL, H. Melih – **The World Wide Web: A New Medium for Statistical Data Dissemination.** Turkey: State Institute of Statistics, 1997.

ASSO Developers, Partners – **User Manual for SODAS 2 Software.** Analysis System of Symbolic Official Data: Edited by FUNDP, 2004. [Consult. 15 Mar. 2005]. Disponível na WWW: <URL: <http://www.ceremade.dauphine.fr/%7Etuati/sodas-pagegarde.htm>>.

BANCO DE PORTUGAL – **Lei Orgânica.** Portugal, 2004.

BAPTISTA, Paulo – **A Inovação nos Produtos, Processos e Organizações.** Sociedade Portuguesa da Inovação, Lisboa: Principia Editora, 1999.

BELKIN, N. J. – **Towards a Definition of Information for Informatics.** In Horsnell, V, (ed), Informatics 2: Proceedings of a Conference Held by the ASLIB Co-ordinate Indexing Group, 25-27 March 1974, New College Oxford, London: Aslib, 53, 1975.

BELKIN, Nicholas J.; MARCHETTI, Pier Giorgio, COOL, C.ollen. – **BRAQUE: Design of an Interface to Support User Interaction in Information Retrieval.** Information Processing and Management, 1993, 29 (3):325-344.

BERINSTEIN, Paula – **The Numbers Game: The Top 10 Sources for Statistics.** Online ABI/INFORM Global, 22, 2, 1998. p. 61-65. [Consult. 20 Set. 2004]. Disponível na WWW: <URL: <http://www.onlinemag.net/OL1998/berinstein3.html>>.

BILLARD, L.; DIDAY, E. – **Symbolic Data Analysis: Definition and Examples**. SODAS Publications, 2002. [Consult. 15 Mar. 2005].

Disponível na WWW: <URL: <http://www.ceremade.dauphine.fr/%7Eetuati/biblio.htm>>.

BOCK, H. H.; DIDAY, E. – **Analysis of Symbolic Data: Exploratory methods for extracting statistical information from complex data**. Berlin: Springer, 2000.

BRAZEAL, Deborah V.; HERBERT, Theodore T. – “**The Genesis of Entrepreneurship: Change, Innovation, and Creativity**”. Estadistic State Polytechnic University [Consult. 10 Set. 2004]. Disponível na WWW:

<URL: <http://www.sbaer.uca.edu/Research/1996/USASBE/96usa283.txt>>.

BRITO, Paula – **Hierarchical and Pyramidal Clustering for Symbolic Data**. J. Jpn. Soc. Comp. Statist., 15, 2002. p.1-14

CARNEIRO, Alberto – **Inovação, Estratégia e Competitividade**. Textos de Gestão, 1.^a ed. Lisboa: Texto Editora, 1995.

CEAPARU, Irina – **Finding Governmental Statistical Data on the Web: A Case Study of Fedstats**. IT&Society, Vol.1, 3, 2003. p. 1-17. [Consult. 14 Set. 2004].

Disponível na WWW:

<URL: <http://www.stanford.edu/group/sigss/itandsociety/v01i03.html>>.

CHOO, Chun Wei – **The Knowing Organization: How Organizations Use Information to Construct Meaning, Create Knowledge, and Make Decisions**. New York: Oxford University Press, 1998.

CHOO, Chun Wei; AUSTER, Ethel, eds. – **Environmental Scanning: Acquisition and Use of Information by Managers**. Edited by M. E. Williams. Vol. 28, Annual Review of Information Science and Technology, Medford, NJ: Learned Information, Inc., 1993.

CHOO, Chun Wei; DETLOR, Brian; TURNBULL, Don – **Information Seeking on the Web: An Integrated Model of Browsing and Searching**. First Monday 5(2). [Consult. 21 Ago. 2004]. Disponível na WWW:

<URL: http://www.firstmonday.dk/issues/issue5_2/choo/index.html>.

CHOO, Chun Wei; DETLOR, Brian; TURNBULL, Don – **Web Work: Information Seeking and Knowledge Work on the World Wide Web**. Dordrecht, Netherlands: Kluwer Academic Publishers, 2000.

COMISSÃO EUROPEIA, **Livro verde sobre a inovação**. Luxemburgo: Gabinete de Publicações Oficiais das Comunidades Europeias, 1996.

CONCEIÇÃO, Pedro; ÁVILA, Patrícia – **A Inovação em Portugal, II Inquérito Comunitário às Actividades de Inovação**. Observatório das Ciências e das Tecnologias, Oeiras: Celta Editora, 2001.

CONSELHO SUPERIOR DE ESTATÍSTICA – **Análise e Reflexão sobre as Normas Actuais do Instituto do Segredo Estatístico**. N.º 2, Lisboa: Instituto Nacional de Estatística, 1999.

COOKE, Alison – **A Guide to Finding Quality information on the Internet – Selection and Evaluation Strategies**. London: Library Association Publishing, 1999.

CRISTAS, Assunção – **O Conselho Superior de Estatística, O Conceito de Segredo Estatístico e a Aplicação Prática da Legislação referente à Libertação de Dados sujeitos a Segredo Estatístico**. Seminário Princípio do Segredo Estatístico, Lisboa: Instituto Nacional de Estatística, 2005.

DANIELS, N. Caroline – **Estratégias Empresariais e Tecnologias da Informação**. Lisboa: Biblioteca de Economia e Gestão, Editorial Caminho, 1997

DERVIN, Brenda – **From the Mind's Eye of the 'User': The Sense-Making Qualitative-Quantitative Methodology**. In *Qualitative Research in Information Management* editado por J.D. Glazier e R.R. Powell, Englewood, Libraries Unlimited, 1992.

DIÁRIO DA REPÚBLICA – **Lei de Bases do Sistema Estatístico Nacional**. I Série, N.º 88, Lei 6/89 de 15 de Abril, art. 5º.

DIAS, Carlos S. Afonso – **A Componente Geográfica nas Estatísticas Oficiais**. Lisboa: Instituto Nacional de Estatística, 2002.

DIAS, Carlos Sebastião Afonso – **Estratégias de Difusão: Que meios para que público?**. Revista de Estatística, 2º Quadrimestre, Vol.2, Instituto Nacional de Estatística, 1999.

DIDAY, Edwin – **An Introduction to Symbolic Data Analysis and The SODAS Software**. The Electronic Journal of Symbolic Data Analysis, Vol.0, N. 0, 2002.

DITKOFF, Mitchell – **Qualities of an Innovator**. Bakersfield, CA: InnovationNetwork, 2004. [Consult. 10 Set. 2004].
Disponível na WWW: <URL: <http://www.thinksmart.com/articles/qualities.html>>.

DOMINGO-FERRER, Josep, MATEO-SANZ, JOSEP M. – **On Resampling for Statistical Confidentiality in Contingency Tables**. Tarragona, Catalonia, Universitat Rovira I Virgili, 1999.

DRUCKER, Peter F. – **Inovação e Gestão, Uma Nova Concepção de Estratégia de Empresa**. 3.ª ed. Lisboa: Biblioteca de Gestão Moderna, Editorial Presença, 1989.

ELLIS, David – **A Behavioural Model for Information Retrieval System Design**. Journal of Information Science 15, nº4/5, 1989. p. 237-247

ELLIS, David – **Modelling the Information Seeking Patterns of Engineers and Research Scientists in an Industrial Environment**. Journal of Documentation 53 (4), 1997. p. 384-403

ELLIS, David – **New Horizons in Information Retrieval**. London: Library Association, 1991.

FERREIRA, Maria João; TAVARES, Isabel – **Notas sobre a História da Estatística**. Dossier VI, 2002. [Consult. 14 Set. 2004].

Disponível na WWW:

<URL: <http://alea-estp.ine.pt/html/statofic/html/dossier/html/dossier.html>>.

FIENBERG, Stephen E. – **Confidentiality and Disclosure Limitation Methodology: Challenges for National Statistics and Statistical Research**. Pittsburgh: Department of Statistics Technical Reports, Carnegie Mellon University, 1997.

FREIRE, Adriano – **Inovação, Novos Produtos, Serviços e Negócios para Portugal**. Lisboa: Editorial Verbo, 2000. p. 389-399.

GAMA, Ana – **A Realização de Estudos nos Institutos de Estatística**. Portugal, Revista de Estatística, 2º Quadrimestre, nº 2, Instituto Nacional de Estatística, 1996.

GARRIDO, Patrícia C. – **Applications of Symbolic Objects in Official Statistics**. Vitoria-Gasteiz, País Basco, EUSTAT – Instituto Vasco de Estadística, 2000.

GHIGLIONE, Rodolphe; MATALON, Benjamin – **O Inquérito, Teoria e Prática**. 2.^a ed. , Oeiras: Celta Editora, 1993.

GIOVANNINI, Enrico – **Damned lies and Statistics: helping numbers make sense**. Organisation for Economic Cooperation and Development. The OECD Observer, ABI/INFORM Global, nº 237, 2003. p. 51-52

GRACIOSO, Luciana de Souza – **Disseminação de Informações Estatísticas no Brasil: Práticas e Políticas das Agências Estaduais de Estatística**. Brasília: Ci. Inf., v. 32, 2, 2003. p. 69-76.

HAMEL, Gary; PRAHALAD, C. K. – **Competindo pelo Futuro: Estratégias Inovadoras para Obter o Controle do seu Sector e Criar os Mercados de Amanhã**. Harvard Business School Press, Rio de Janeiro: Editora Campus, 1995.

HILL, Andrew; HILL, Manuela M. – **Investigação por Questionário**. 2.^a ed. Lisboa: Edições Sílabo, 2002.

HINCHLIFFE, Lisa; DUKE, Lynda – **Finding Statistics: Tutorial and Worksheet**. Illinois State University: Milner Library, 2001. [Consult. 20 Set. 2004].
Disponível na WWW: <URL: <http://www.mlb.ilstu.edu/learn/stat/home.htm>>.

INE DESTAQUE, Informação à Comunicação Social – **Inquérito à Utilização das Tecnologias e da Comunicação nas Empresas 2003**. Sociedade da Informação e do Conhecimento. Lisboa: INE, Fevereiro 2004.

INNOVATIONNETWORK – **Innovation Proverbs Poster**. Bakersfield, CA: InnovationNetwork, 2003. [Consult. 10 Set. 2004].
Disponível na WWW: <URL: <http://thinksmart.com/library/index.html>>.

JAVEAU, C. – **L'enquête par questionnaire**. 2.^a ed. Bruxelles : Éditions de l'Université de Bruxelles, 1978.

JCASÓ, Peter – **Finding Statistical Information on CD ?**. Information Today, 15, 10 ABI/INFORM Global, 1998. p. 46-47

JOHANNESSEN, Jon-Arild; OLSEN, Bjorn; LUMPKIN, G.T. – **Innovation as newness: what is new, how new, and new whom?** European Journal of Innovation Management 4 (1), 2001. p.20-31

JOHNSTON, Donald J. – **Statistics, Knowledge and Progress**. The OCDE Observer, n.º 246-247, 2005. [Consult. 25 Mar. 2005].
Disponível na WWW: <URL: http://www.oecdobserver.org/news/printpage.php/aid/1514/Statistics_knowledge_and_progress.html>.

KUHLTHAU, Carol C. – **Inside the Search Process: Information Seeking from the User's Perspective**. Journal of the American Society of Information Science 42 (5), 1991. p. 361-371

KUHLTHAU, Carol C. – **Seeking Meaning: A Process Approach to Library and Information Services**. Norwood, NJ: Ablex Publishing Corporation, 1993.

LARGE, Andrew; TEDD, Lucy A.; HARTLEY, R.J. – **Information Seeking in the Online Age, Principles and Practice**. UK: Bowker-Saur, 1999.

LEINER, Barry M., [et al.] – **A Brief History of the Internet**. Internet Society, 2003. [Consult. 20 Ago. 2004].
Disponível na WWW: <URL: <http://www.isoc.org/internet/history/brief.shtml>>.

MACMULLIN, Susan E.; TAYLOR, Robert S. – **Problem Dimensions and Information Traits**. Information Society 3, n.º 1, 1984. p. 91-111

MARCELO, Carlos – **Application of Symbolic Data Analysis to Official Statistical Data**. Lisboa: Workshop on Applications of Symbolic Data Analysis, 2004.

MARCELO, Carlos; BRITO, Paula – **Symbolic Data Analysis for Official Statistics, The SODAS Software.**, Aveiro: X Jornadas de Classificação e Análise de Dados, JOCLAD 2003.

MARCELO, Carlos; SILVA, Ana A.; RODRIGUES, Duarte – **Especificidades Regionais do Emprego em Portugal: O Contributo da Análise Multivariada de Objectos Simbólicos.** Portugal, Revista de Estudos Regionais de Lisboa e Vale do Tejo, nº 2, Instituto Nacional de Estatística, 2001.

MARCHIONINI, Gary – **Information Seeking in Electronic Environments.** UK Cambridge: Cambridge University Press, 1995.

MAROCO, João – **Análise Estatística Com Utilização do SPSS.** 2.^a ed. Lisboa: Edições Sílabo, 2003.

MAS, Marta, [et al.] – **Symbolic Characterisation of Internet Users in the Basque Country.** País Basco, EUSTAT-Instituto Vasco de Estadística, 2002.

MAS, Marta; OLAETA, Haritz – **Symbolic Analysis of the Time Use in the Basque Country.** País Basco, EUSTAT-Instituto Vasco de Estadística, 2002.

MORINEAU, A. – **SODAS, Symbolic Official Data Analysis System.** EUROSTAT, Research in Official Statistics. [Consult. 20 Mar. 2005]
Disponível em:
<URL:<http://europa.eu.int/en/comm/eurostat/research/dosis/sodas/sld001.htm>>.

NANOPOULOS, Ph.; KING, John – **Important Issues on Statistical Confidentiality Methods.** Luxemburgo: Eurostat, European Commission, 2002.

OLIVEIRA, Jorge – **Gestão da Inovação.** Sociedade Portuguesa da Inovação, Lisboa: Principia Editora, 1999.

PARDAL, L.; CORREIA, E. – **Métodos e Técnicas de Investigação Social.** Porto: Areal Editores, 1995.

PARISIAN, Monique; PODEHL, Martin – **Effective Presentation of Statistics in Electronic Dissemination.** Canada: Statistics Canada, 1997

PEREIRA, Alexandre; POUPA, Carlos – **Como Escrever uma Tese, monografia ou livro científico usando o Word.** 2.^a ed., Lisboa: Edições Sílabo, 2003.

PEREIRA, H. Caeiro – **Banco de Portugal do Passado ao Presente.** Lisboa: Serviço de Tipografia do Banco de Portugal, 1985.

PESSOA, Fernando – **A Essência do Comércio.** Lisboa: Revista de Comércio e Contabilidade, 1, 1926.

PESTANA, Maria Helena; GAGEIRO, João Nunes – **Análise de Dados Para Ciências Sociais, A Complementaridade do SPSS**. 3.^a ed., Lisboa: Edições Sílabo, 2003.

PORTER, Michael – **The Competitive Advantage of Nations**. The Free Press, 1990.

SACHS, L. – **Applied Statistics, a Handbook of Techniques**. 2.^a ed. New York: Springer-Verlag.

SARAIVA, Pedro M.; D'OREY, João - **Inovação e Qualidade**. Sociedade Portuguesa da Inovação, Lisboa: Principia Editora, 1999.

SOUSA, Fernando – **História da Estatística em Portugal**. Lisboa: Instituto Nacional de Estatística, 1995.

SOUSA, João Luís – **Presidente do INE (Paulo Gomes) considera que as Empresas devem utilizar mais Informação Estatística**. Vida Económica, Ed. 18 Julho de 2003. [Consult. 3 Mar. 2004].
Disponível na WWW: <URL: <http://www.centroatl.pt/ve/jul2003/ve1012emp.html>>.

STATISTICS CANADA - **CANSIM User Guide**. Revised November 5. Ottawa: Statistics Canada, 2004.

TAYLOR, Robert S. - **Information Use Environments, In Progress in Communication Science**, editado por B. Dervin e M.J. Voigt, Norwood, NJ: Ablex Publishing Corporation, 1991.

TESKEY, F. N. - **User Models and World Models for Data, Information and Knowledge**. Information Processing and Management, 25 (1), 8, 1989.

TURNBULL, Don - **Augmenting Information Seeking on the World Wide Web Using Collaborative Filtering Techniques**, 1995. [Consult. 21 Ago. 2004]. Disponível na WWW: <URL: <http://www.ischool.utexas.edu/~donturn/research/augmentis.html>>.

UTTERBACK, James M. – **Dominando a Dinâmica da Inovação**. Brasil: QualityMark Editora, 1996. p. 1-24.

WILSON, Tom - **Information Behavior: An Interdisciplinary Perspective**. Information Processing & Management, 33, n.º 4, 1997. p. 551-572.

Referências Electrónicas

Adobe System. [consultado em 15 Fev.2005].

URL: <http://www.adobe.com>

ALEA – Acção Local de Estatística Aplicada. [Consult. 2 Fev. 2005]

URL: <http://alea-estp.ine.pt/>

American Demographics. [Consult. 17 Set. 2004]

URL: <http://www.demographics.com>

American Marketing Association. [Consult. 7 Fev. 2005]

URL: <http://www.marketingpower.com>

American Statistical Association. [Consult. 20 Set. 2004]

URL: <http://www.amstat.org>

ASSO - Analysis System of Symbolic Official Data [Consult. 10 Mar. 2005]

URL: <http://www.assoproject.be>

Associação Industrial Portuguesa. [Consult. 7 Fev. 2005]

URL: <http://www.aip.pt>

Banco de Portugal. [Consult. 20 Set. 2004]

URL: <http://www.bportugal.pt>

Base de Conceitos Estatísticos INE. [Consult. 8 Fev. 2005]

URL: <http://conceitos.ine.pt/>

Base de Fontes Estatísticas INE. [Consult. 8 Fev. 2005]

URL: <http://gefstat.ine.pt/>

CANSIM – Canadian Sócio-economic Information Management System. [Consult. 3 Jan. 2005]

URL: <http://cansim2.statcan.ca>

Conceitos Estatísticos do INE. [Consult. 2 Fev. 2005].

URL: <http://conceitos.ine.pt>

Depository Services Program – Statistics Canada. [Consult. 3 Jan. 2005]

URL: <http://dsp-psd.pwgsc.gc.ca/index-e.html>

DGEP – Direcção-Geral de Estudos e Previsão. [Consult. 20 Set. 2004]

URL: <http://www.dgep.pt/menuprinc.html>

Encarta Online. [Consult. 17 Set. 2004]

URL: <http://encarta.msn.com/>

Encyclopedia.com. [Consult. 31 Ago. 2004]

URL: www.encyclopedia.com

EUROSTAT – Statistical Office of the European Communities. [Consult. 31 Out. 2004]

URL: <http://europa.eu.int/comm/eurostat>

EUSTAT – Instituto Vasco de Estadística. [Consult. 15 Mar. 2005]

URL: <http://www.eustat.es/>

FEDSTATS. [Consult. 17 Set. 2004]

URL: <http://www.fedstats.gov>

First Monday – Peer-Reviewed Journal on the Internet. [Consult. 20 Ago. 2004]

URL: <http://www.firstmonday.dk>

Fontes Estatísticas do INE. [Consult. 2 Fev. 2005].

URL: <http://conceitos.ine.pt>

Forrester Research. [Consult. 7 Fev. 2005]

URL: <http://www.forrester.com>

Fulcrum Analytics. [Consult. 7 Fev. 2005]

URL: <http://www.cyberdialogue.com>

Glossário *Statistics Canada*. [Consult. 30 Out. 2004]

URL: <http://www.statcan.ca/english/edu/power/glossary/gloss.htm>

HighBeam Library Research. [Consult. 17 Set. 2004]

URL: <http://www.highbeam.com/library/index.asp>

Innovation Network. [Consult. 10 Set. 2004]

URL: <http://www.thinksmart.com/index.html>

Instituto Nacional de Estatística. [Consult. 20 Ago. 2004]

URL: <http://www.ine.pt>

Internet Society (ISOC). [Consult. 20 Ago. 2004]

URL: <http://www.isoc.org>

IT&Society – A Web Journal Studying How Technology Affects Society. [Consult. 15 Set. 2004]

URL: <http://www.ITandSociety.org>

Journal of Symbolic Data Analysis [Consult. 10 Mar. 2005]

URL: <http://www.jsda.unina2.it>

Marktest [Consult. 7 Fev. 2005]

URL: <http://www.marktest.pt>

Ministério da Ciência, Inovação e Ensino Superior. [Consult. 20 Set. 2004]

URL: <http://www.mces.gov.pt/>

Ministério da Justiça. [Consult. 20 Set. 2004]

URL: <http://www.mj.gov.pt>

Ministério das Finanças e da Administração Pública. [Consult. 20 Set. 2004]

URL: <http://www.min-financas.pt>

OCDE – Organização para a Cooperação e o Desenvolvimento Económico. [Consult. 8 Fev. 2005]

URL: <http://www.oecd.org/home/>

OCES – Observatório da Ciência e do Ensino Superior. [Consult. 20 Set. 2004]

URL: <http://www.oces.mces.pt/>

ONS – Office for National Statistics. [Consult. 15 Mar. 2005]

URL: <http://www.statistics.gov.uk/>

SciELO Brasil – Scientific Electronic Library Online. [Consult. 30 Maio 2004]

URL: <http://www.scielo.br>

SODAS - *Symbolic Official Data Analysis System* [Consult. 10 Mar. 2005]

URL: <http://www.ceremade.dauphine.fr/%7Eetuati/sodas-pagegarde.htm>

Statistical Abstract of the U.S. [Consult. 17 Set. 2004]

URL: http://www.census.gov/stat_abstract/

Statistics Canada. [Consult. 30 Out. 2004]

URL: <http://www.statcan.ca/>

Statistics Netherlands. [Consult. 29 Dez. 2004]

URL: <http://www.cbs.nl/en/>

StatLine. [Consult. 29 Dez. 2004]

URL: <http://statline.cbs.nl/StatWeb>

The CIA's World Factbook. [Consult. 17 Set. 2004]

URL: <http://www.cia.gov/cia/publications/factbook/>

The Literature Network [Consult. 15/04/2005]

URL: <http://www.online-literature.com>

The OECD Observer [Consult. 25 Mar. 2005]

URL: <http://www.oecdobserver.org>

U.S. Bureau of Labor Statistics. [Consult. 17 Set. 2004]

URL: <http://www.stats.bls.gov>

U.S. Bureau of the Census. [Consult. 17 Set. 2004]

URL: <http://www.census.gov>

UMIC – Unidade de Missão Inovação e Conhecimento. [Consult. 14 Set. 2004]

URL: <http://www.umic.gov.pt/UMIC>



**Isabel Maria da Silva
Machado**

**A Internet como Factor de Inovação no Acesso a
Informação Estatística**

ANEXOS



**Universidade de
Aveiro
2005**

Departamento de Economia, Gestão e
Engenharia Industrial

**Isabel Maria da Silva
Machado**

**A Internet como Factor de Inovação no Acesso a
Informação Estatística**

**ANEXO I – Critérios de Avaliação de Bases de Dados
Disponíveis *Online* (Cooke, 1999)**

CrITÉrios de Avaliação de Bases de Dados DisponÍveis *Online* **(Cooke, 1999)**

- ☒ What is the intended purpose, coverage and audience of the database? Is this information available from any introductory information or help files?
- ☒ What is the coverage of the database? What subject areas and materials are covered? Is the database comprehensive within a particular area? What is the retrospective coverage of the database? Is information available regarding the material which is included in the database? Are there links to further sources of information?
- ☒ Are different versions of the same database available? Are there any differences in coverage?
- ☒ How much information is provided in each record of the database? Is the amount of information useful? Is the amount of information sufficient for the needs of the user?
- ☒ For bibliographical databases: do they consist of references only, or are abstracts available? For what percentage of the database are abstracts provided? Have the abstracts been truncated by word length?
- ☒ Is it possible to identify the authors responsible for the records in the database? What are their knowledge and expertise?
- ☒ What is the reputation of the database? Is the database included in any guides to reference works, such as *Walford's*?
- ☒ Are there any typographical or spelling errors? Are there any errors in bibliographical citations? Are there any quality-control procedures?
- ☒ Is the database current and well-maintained? How frequently is the database updated? Is there a time delay between the publication of materials and their appearance in the database?
- ☒ Is the information well-presented and arranged? What searching and browsing facilities are available? Are any searching and browsing facilities useful, effective and easy to use? What outputting and downloading options are available? Is it easy to output and downloading data from the database? Can data be exported into another package? Is this easy to do?
- ☒ In relation to bibliographical databases: is it possible to search by author, title or subject keyword? Can searches be limited by publication type and data range? Are there any additional searching or browsing features? Are the search and browse facilities effective easy to use and useful?

- ☒ Is it possible to order material directly from the database? Is contact information available?
- ☒ Is the database easily accessible? Are there any restrictions to access? Is there a charge to access the database or to download records?
- ☒ Is the database easy to use, and are there any user support facilities?
- ☒ How does the database compare with other similar databases?
- ☒ What overall impression of the quality of the database is created after examining it?



**Isabel Maria da Silva
Machado**

**A Internet como Factor de Inovação no Acesso a
Informação Estatística**

ANEXO II – Questionário

QUESTIONÁRIO



Venho pedir a sua colaboração num trabalho que tem como objectivo aprofundar o conhecimento sobre o grau de utilização do serviço INFOLINE como factor de inovação no acesso a Informação Estatística.

O anonimato será completamente assegurado, pelo que agradeço que responda com a maior sinceridade, pois assim contribuirá para uma melhor investigação na qual este trabalho se insere.

Universidade de Aveiro – Departamento de Economia, Gestão e Engenharia Industrial

1. Qual o seu Sector de Actividade? _____

2. Utiliza o Serviço INFOLINE do INE? Sim ☐ Não ☐

Se respondeu Sim, por favor continue o questionário. Se respondeu Não passe para a **questão 9**.

3. Como teve conhecimento do Serviço INFOLINE?

<input type="checkbox"/> Motor de Pesquisa	<input type="checkbox"/> Link a partir de outra página
<input type="checkbox"/> Comunicação Social	<input type="checkbox"/> E-mail
<input type="checkbox"/> Publicações/Panfletos	<input type="checkbox"/> Através do INE
<input type="checkbox"/> Informação Oral	<input type="checkbox"/> Outro. Por favor, Indique: _____

4. Com que frequência o utiliza?

<input type="checkbox"/> Diariamente	<input type="checkbox"/> Pelo menos 1 vez por mês
<input type="checkbox"/> Pelo menos 1 vez por semana	<input type="checkbox"/> Outra. Por favor, especifique: _____

5. Que tipo de Informação acede no Serviço **INFOLINE**?

<input type="checkbox"/> Destaques	<input type="checkbox"/> Unidades Territoriais
<input type="checkbox"/> Indicadores Estatísticos	<input type="checkbox"/> Informação Rápida
<input type="checkbox"/> Quadros Estatísticos	<input type="checkbox"/> Retratos Territoriais
<input type="checkbox"/> Séries Cronológicas	<input type="checkbox"/> Estudos
<input type="checkbox"/> Novidades	

6. Assinale os temas que mais consulta:

<input type="checkbox"/> Estatísticas Gerais	<input type="checkbox"/> Agricultura, Produção Animal, Silvicultura e Pescas
<input type="checkbox"/> Território e Ambiente	<input type="checkbox"/> Indústria, Construção e Energia
<input type="checkbox"/> População e Condições Sociais	<input type="checkbox"/> Comércio Interno, Turismo e Outros Serviços
<input type="checkbox"/> Economia e Finanças	<input type="checkbox"/> Sociedade da Informação e Conhecimento
<input type="checkbox"/> Comércio Externo	<input type="checkbox"/> Diversos

7. Que apreciação faz do Serviço **INFOLINE** segundo os seguintes aspectos?

	<i>Muito Mau</i>				<i>Muito Bom</i>	
	1	2	3	4	5	6
Organização da Informação	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Quantidade e Diversidade de Informação	((((((
Interactividade	((((((
Design	((((((
Sistema de Navegação	((((((
Conteúdos	((((((
Actualização de Conteúdos	((((((
Indique, se for o caso, outro aspecto do INFOLINE que não tenha sido aqui considerado:	((((((

8. Em geral, encontra no INFOLINE a Informação que procura? Sim (Não (

Se respondeu Não. Indique, então, que tipo de Informação gostaria de encontrar no INFOLINE:

9. Acede a outros sites de informação estatística/divulgação de dados estatísticos? Sim (Não ☐

Se respondeu Sim. Indique, por favor, qual ou quais sites acede:

Se respondeu Não. Indique, por favor, o motivo para não utilizar:

Obrigado pela sua Colaboração!

Isabel Machado
Mestrado em Gestão de Ciência, Tecnologia e Inovação